Uncovering the Mechanism of Shenlian Capsules in the Treatment of Non-small Cell Lung Cancer by System Pharmacology

Wanjin Fu

Department of Clinical Pharmacology, The Second Hospital of Anhui Medical University, Hefei, 230601, China

Ajuan Gong

Department of Endocrinology, The Second Hospital of Anhui Medical University, Hefei, 230601, China

Bin Yu

Department of Clinical Pharmacology, The Second Hospital of Anhui Medical University, Hefei, 230601, China

Xin Wei

Department of Clinical Pharmacology, The Second Hospital of Anhui Medical University, Hefei, 230601, China

Yong Chen

Department of Clinical Pharmacology, The Second Hospital of Anhui Medical University, Hefei, 230601, China

Renpeng Zhou

Department of Clinical Pharmacology, The Second Hospital of Anhui Medical University, Hefei, 230601, China

Chao Lu

Department of Clinical Pharmacology, The Second Hospital of Anhui Medical University, Hefei, 230601, China

Wei Hu (✉ 1063775796@qq.com)

Department of Clinical Pharmacology, The Second Hospital of Anhui Medical University, Hefei, 230601, China <https://orcid.org/0000-0002-8819-4641>

Research

Keywords: System pharmacology, Shenlian Capsule, non-small cell lung cancer, bioinformatics, clusters, signaling pathway

Posted Date: August 5th, 2020

Abstract

Background: Shenlian Capsule is a Chinese medicine compound approved by the China Food and Drug Administration (CFDA) for the treatment of advanced non-small cell lung cancer(NSCLC). However, due to its complex constituents, cause its effective active compounds and main action mechanisms for treating diseases are still not fully clear. The purpose of this work is to explore the active ingredients and mechanisms of Shenlian Capsule treatment NSCLC through a system pharmacological approach.

Methods: First, a database of Shenlian Capsule chemical composition was constructed by retrieving Chinese herbal medicine data. Absorption, distribution, metabolism and excretion (ADME) methods were used to screen potential active compounds. Predict active compound targets with a large-scale molecular network target prediction technology. Clustering of active compounds obtained through cluster analysis by MECODE plug-in, each cluster obtains the main pathways through enrichment analysis method. To get all targets related to survival in NSCLC, the survival related targets were intersected with the compounds target (C-T) network to get the survival related targets for Shenlian Capsule. Finally, a batch molecular docking technique was used to verify the effect of active compounds of Shenlian Capsule on survival-related targets.

Results: The Shenlian Capsule C-T network was constructed with 117 potential compounds and 47 targets. Treatment of NSCLC with Shenlian Capsule through enrichment analysis may involve multiple pathways such as anti-inflammatory, immune regulation, regulation of cell cycle, apoptosis, antiviral, cell hypoxia, angiogenesis and so on. Shenlian Capsule has eight survival-related genes in non-small cell lung adenocarcinoma (LUAD) and two survival-related genes in non-small cell squamous cell lung carcinoma (LUSC). It is known through molecular docking that the active compound has lower energy after conformation with survival-related genes, and has lower binding energy and stable binding.

Conclusion: In this study, the potential compounds of active compounds in Shenlian Capsule were first predicted using network pharmacology technology. Through the enrichment analysis, the main pathways of the role of Shenlian Capsule were outcropped. Secondly, by combining bioinformatics and network pharmacology, Shenlian Capsule can be regulated to target survival-related targets. Finally, the molecular docking technique shows the relationship between active compounds and survival-related targets after docking. This work provided a scientific basis for the clinical role of Shenlian Capsule in the treatment of NSCLC, provided a research basis for further clarifying the active ingredients and mechanism of Shenlian Capsule in the treatment of NSCLC.

Introduction

Lung cancer is the leading cause of cancer-related deaths not only in China but throughout the world. In recent years, the incidence of lung cancer has risen sharply due to environmental factors such as smoking[1, 2]. Non-small cell lung cancer (NSCLC) is the main type of lung cancer, accounting for more than 85% of the incidence of lung cancer. The early clinical symptoms of NSCLC are not obvious, it is

found that the diagnosis of symptoms is usually in the middle or late stage of cancer which leads to a higher mortality rate[3]. The 5-year survival rate is only 16.8% after the initial diagnosis of the disease in the United States[4]. At present, the main measures for the treatment of NSCLC include chemotherapy and radiotherapy in addition to surgery. Chemotherapy and radiation therapy can improve patients' quality of life and prolong life, but it generally cannot cure NSCLC. In addition to killing tumor cells, chemotherapy also has great damage to normal human cells. For example, chemotherapy can lead to inhibition of the bone marrow hematopoietic system, leading to a decline in white blood cells and platelets. Radiotherapy for lung cancer can lead to complications such as radiation pneumonitis, radiation esophagitis, radiation pulmonary fibrosis and radiation-induced myelitis[5]. Lung cancer has become a common challenge for public health systems around the world.

Traditional Chinese medicine (TCM) is often a synergistic effect of multi-component, multi-target and its components to treat diseases. As an important part of supplements and alternative medicines, Chinese herbal medicines including Shenlian Capsule are also used clinically to treat NSCLC. Shenlian Capsule is a drug approved by the CFDA for the treatment of patients with intermediate and advanced lung cancer. The Shenlian Capsule consists of BaiBianDou(BBD), BanZhiLian(BZL), BuGuZhi(BGZ), DanShen(DS), EZhu(EZ), FangJi(FJ), KuShen(KS), KuXingRen(KXR), SanLeng(SL), ShanDouGen(SDG) and WuMei(WM) eleven kinds of Chinese herbal medicine (Fig. 1A)[6]. Previous clinical trials have shown that erlotinib combined with Shenlian Capsule in the treatment of NSCLC in 40 patients with NSCLC has a definite therapeutic effect and mild adverse reactions[7]. Due to economic and technical reasons, pharmacological research on the mechanism of action of multi-component and multi-target drugs is still not deep[8]. The specific mechanism of action of Shenlian Capsule in the treatment of lung cancer is also not yet clear.

In light of Shenlian Capsule complex composition and diversity of targets, it is very difficult to study the mechanism of action of Shenlian Capsule on diseases using current pharmacological research methods. System pharmacology is a new discipline that has developed rapidly in recent years. It is a multidisciplinary product, including pharmacology, biochemistry and structural biology, genomics, medicine and applied mathematics, computer technology, bioinformatics, etc [9, 10]. System pharmacology technology was used to study the active substances and combinations of Shenlian Capsule in the treatment of NSCLC, and to identify the target of pharmacodynamic components, pharmacodynamic substances and disease targets to provide theoretical basis for the clinical application of Shenlian Capsule.

Materials And Methods

Establishment of chemical composition database of Shenlian Capsule

To construct the chemical composition database of Shenlian Capsule placed the Latin names of eleven

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js

and the Traditional Chinese Medicine Systems

Pharmacology Database to collect the ingredients of compounds of each herb and the pharmacodynamic parameters of each compound [11, 12]. The results of each of the above herbal search were combined to obtain a Shenlian Capsule ingredient database. In addition, psoralen chemical composition was obtained from Anticancer Herbs database of Systems Pharmacology (Cancer HSP) [13]. In this study, 714 unique compounds were obtained in Shenlian Capsule, which BBD, BZL, BGZ, DS, EZ, FJ, KS, KXR, SL, SDG and WM were collected 14, 94, 4, 202, 81, 50, 113, 113, 30, 79 and 40 compounds, respectively (Fig. 1B).

Bioactive compounds screening

Screening potential active compounds by using computer calculations and predicting ADME parameters for each chemical component is faster, more effective and more economical than traditional animal experiments [12]. In this work used oral bioavailability (OB), drug-like (DL), Caco-2 and half-life (HL) to screen potential constituent compounds in the chemical composition database of Shenlian Capsule.

Oral Bioavailability

The OB index is mainly used to predict the relative amount of the compound into the systemic blood after oral absorption. Many active compounds cannot be active because they are insufficiently absorbed into the body after oral absorption [14]. This indicator is a critical indicator of whether a compound is intrinsically active. In this study, the interaction between the compound and metabolism (P450 3A4) and the transporter (P-glycoprotein) were calculated using a tried OBioavail 1.1 model to obtain the OB value of each compound. The OB value is greater than 30 as a screening active compound threshold [15].

Drug-like evaluation

The DL indicator evaluates the compound as a drug possibility by comparing the similarity between the compound and the known drug. By calculating the DL value, it is advantageous to screen out high-activity potential compounds in the early stage of drug development [16]. In this study, a database-dependent model was applied using the Tanimoto coefficient (see Eq. (1)) to get the DL value of each compound in the Shenlian Capsule.

$$T(a, b) = \frac{a * b}{|a|^2 + |b|^2 - a * b} \quad (1)$$

a represents the value of the descriptor obtained by DRAGON software and b represents the average descriptor value of all drugs in the DrugBank database. The DL value is greater than 0.18 as a screening active compound threshold [8].

Caco-2 permeability

The permeability of the drug in the intestinal epithelial cells determines the extent to which the drug is absorbed by the body after oral administration. The rate of penetration of the drug in a single layer of

Caco-2 cells represents the permeability of the drug in the intestinal epithelial cells. The Caco-2 permeation rate values of the individual compounds in the Shenlian Capsule were calculated by calculation on silicon using the VolSurf method[17]. The Caco-2 permeability value is greater than – 0.04 nm/s as a screening active compound threshold[8].

Half-life (HL)

The drug HL index refers to the concentration of drug in the blood or the time it takes to reduce the amount of drug in the body to one-half which is the most important indicator in the ADME model. In clinical drug half-life indicators can guide the rational compatibility of drugs, determine the dosing interval, determine the dose and estimate the amount of drug accumulation in the body. The HL value is greater than 3 h as a screening active compound threshold[8].

OB indicators screened 340 compounds, DL indicators screened 397 compounds, Caco-2 screened 479 compounds and HL indicators screened 154 compounds(Fig. 1C). Finally, using Venn diagrams software to intersection the results of the above four indicators, a total of 143 potential active compounds were acquired (Fig. 1D).

Prediction of targets of Shenlian Capsule

In this study, the potential active compound InChKey number screened above was placed in the Traditional Chinese Medicine Systems Pharmacology Database to obtain a compound to predict the relationship between the compound and the targets. The database uses a large-scale molecular network target prediction technology, mainly using the WES method and SysDT model to obtain the relationship between compounds and targets[14]. Since the targets obtained by this database include multiple species targets and non-standard gene names for gene names, this is not conducive to clearly showing the relationship between compounds and targets. We acquired the target name from the above database and put it into the Uniport database to select the standard gene name of the species only Homo sapiens[8].

NSCLC-related targets

NSCLC-related target data were obtained from four sources: (i)Online Mendelian Inheritance in Man (OMIM), (ii) Therapeutic Target Database (TTD), (iii) Uniprot, (iv) DIGSEE. Enter the keyword " non-small cell lung cancer " in the above database to search for a disease-related gene of the Homo sapiens. Targets related to NSCLC disease were obtained from TTD, OMIM, DIGSEE and Uniport databases, respectively, with 24, 60, 91, 74 targets. The above results were removed from the repeated targets and a total of 226 NSCLC-related targets were obtained (Fig. 2B).

Construction of Shenlian Capsule compound-disease-target network

In order to better explore and analysis the main mechanism of Shenlian Capsule in the treatment of NSCLC, open source software Cytoscape(version: 3.7.1)[18] was first used to construct the Shenlian Capsule Compound-Protein Target (C-T) network, disease-target (D-T) network. The network merge tool in Cytoscape software was used to intersect the above two networks to build a Shenlian Capsule compound-disease-target (C-D-T) network.

Construction of disease protein-protein interaction (PPI) network

PPI networks are composed of individual proteins that interact with each other to participate in all aspects of life processes such as biosignal delivery, gene expression regulation, energy metabolism, material metabolism and cell cycle regulation. This work used the Bisogenet 3.0 plug-in[19] in Cytoscape to visualize the PPI network of NSCLC-associated target. Bisogenet can automatically obtain protein-protein relationship data from six databases and automatically plot PPI networks in Cytoscape software. The six database names are Database of Interacting Proteins (DIP), Biological General Repository for Interaction Datasets (BioGRID), Human Protein Reference Database (HPRD), IntAct Molecular Interaction Database (IntAct), Molecular INTERaction Database (MINT) and Biomolecular Interaction Network Database (BIND). In this study, the "Method" attribute of Bisogenet has been set to "output nodes one" when acquiring the interaction relationship of each node of the PPI network.

PPI network core targets evaluation

In order to obtain the main mechanism of Shenlian Capsule in the treatment of NSCLC, the intersection of Shenlian Capsule action target network and NSCLC-related targets was first obtained to obtain C-D-T PPI network. The C-D-T PPI network was screened with " $\text{Degree(DC)} \geq 2 \times \text{median DC}$ " as the initial screening condition, the CytoNCA version 2.1.6 [20] plug-in was applied to perform topological analysis on the above results. Topological indicators were used to screen the core targets of the primary screening network, including topological closeness centrality(CC), degree centrality (DC), eigenvector centrality (EC), betweenness centrality (BC), largest average local connectivity (LAC) and network centrality (NC)[21]. Targets screened by topological indicators are the core targets of the C-D-T PPI network.

Cluster analysis

Some target proteins in cell biological activity are closely related and have the same or similar functions, so these target proteins can be considered as one cluster. Proteins in the same cluster have synergistic effects on certain functions of the body. As an important analysis method, clustering analysis can clearly show the interaction rules of nodes in PPI networks[22]. The MCODE plugin in Cytoscape was used for cluster analysis of PPI networks because the higher stability of the module network generated by the molecular complex detection algorithm has been confirmed[23].

Enrichment analysis of GO and KEGG pathway

The four subnets obtained by cluster analysis of the core targets and the protein targets of the C-D-T

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js P), cell component, and molecular functions

(MF)) enrichment analysis and KEGG pathway analysis. Enrichment analysis and KEGG pathway data acquisition mainly use Database for Annotation Visualization and Integrated Discovery (DAVID), KEGG pathway enrichment analysis bubble diagram is mainly produced using the omicshare platform. Enrichment analysis results the study screened $p < 0.01$ and excluded unrelated signaling pathway[24].

Effect of Shenlian Capsule on survival related genes in NSCLC

Data Preparation

The TCGA database was used to download NSCLC gene expression data and patient-related clinical data. Gene expression data download parameters are as follows: the case file parameters are selected as (1) Primary Site choose 'Bronchus and lung', (2) Program choose 'TCGA', (3) Project selects 'TCGA-LUAD' and 'TCGA-LUSC' respectively. Files parameter selection is (1) Data Category was selected as transcriptome profiling, (2) Data Type was selected as Gene Expression Quantification, (3) Experimental Strategy was selected as RNA-Seq, (4) Workflow Type was selected as HTSeq-FPKM. After selecting non-small cell lung adenocarcinoma(LUAD) as described above, 515 cases and 594 files were obtained. Non-small cell squamous cell lung carcinoma(LUSC) received 501 cases and 551 files.

The patient-related clinical data download case parameters are consistent with the gene expression data. The file parameters are selected as follows: (1) Data Category was selected as clinical, (2) Data Type was selected as Clinical Supplement, (3) Workflow Type was selected as HTSeq-FPKM, (4) Data Format select 'bcr xml type'. After the above selection, LUAD gets 522 files and LUSC gets 504 files.

Differential gene analysis

Decompress the obtained gene expression data first, and then use perl script to process a matrix file. A total of 594 LUAD matrix files, including 59 normal tissue samples and 535 cancer tissue samples; a total of 551 samples of LUSC matrix files, including 49 normal tissue samples and 502 cancer tissue samples.

Use the perl script file to convert the gene name format of the matrix to the official gene symbol format. The R matrix script was used to process the above matrix file. First, the duplicate gene name sample expression amount was averaged, and the average gene expression amount less than 0.1 data was deleted to obtain a new matrix file. The Wilcox test was used on the matrix obtained above to compare genes with different gene expression levels between normal tissue and cancer tissue, and the aforementioned differential genes were written into a differential gene expression matrix[25]. The differential gene expression matrix filtering conditions are P-value < 0.05 and the tumor tissue expression level is half or two times the normal tissue expression level. Finally, R script files were used to show the volcano maps and heatmaps of the differentially expressed genes.

Survival related gene screening

After downloading clinically relevant data from TCGA and decompressing it, the perl script was used to obtain clinically relevant data files. Delete the samples with a survival time of less than 30 days and no survival status in the above file. The filtered clinical data file and the differential gene expression file are combined to obtain a clinical-differential gene expression file containing the sample name, survival status, survival time and differential gene expression data. The clinical-differential gene expression file was analyzed by Kaplan-Meier (KM) method[26] for each differential gene, the screening conditions were $P < 0.01$ and genes related to the survival of LUCA or LUSC were obtained.

Survival-related genes merge with C-T network

Intersect the survival-related gene files with the C-T network to obtain the survival-related genes acted by the Shenlian Capsule and obtain the active compounds acting on the survival-related genes through the C-T network.

Expression of survival-related genes and survival curve of NSCLC by treated with Shenlian Capsule

The Wilcoxon test were used to compare the difference between the expression of survival-related targets acted by Shenlian Capsule in normal tissue and cancer tissue, the filter condition is $P < 0.01$ [27]. The beeswarm package[28] were used to make differential patterns of survival-related genes in different tissues in R scripts. Survival package was used to make survival curves for survival-related genes. The difference value filter condition for high and low expression groups was $P < 0.01$ and the five-year survival rate difference between high and low expression groups was calculated.

Molecular docking verification

In order to verify whether activated compounds act on survival-related proteins, this study use molecular docking to verify. Use a python script to download the structure files of active compounds that interact with survival-related target proteins in batches, the file format is mol2. Search 3D structures related to survival-related genes in Uniport and download 3D structure PDB format files related to survival-related genes from the RCSB website. Use the AutoDock vina internal python script to batch convert small molecule mol2 format files and protein molecule PDB files into charged PDBQT files. AutoDock Tools version 1.5.6 software was used to set the docking parameter file. AutoDock Vina software was used for batch molecular docking, each active compound molecule and protein receptor was set to produce 20 more stable conformations [29].

Use Discovery Studio version 3.5.0 software[30] to convert the PDBQT file of the docking result into a PDB format file. LigPlot version 1.4.4 software[31] were used to make the two-dimensional structure of the active compound after docking with the target protein and PyMol version 1.6[32] were used to display the three-dimensional structure of the docking result. Finally, the lower energy conformation was selected for the survival-related gene docking results for display.

Results

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js

Shenlian Capsule C-T network analysis

Twenty-six of the 143 potential active compounds obtained by the ADME model were excluded due to they did not predict a target for interaction with humans. The C-T network of the Shenlian Capsule consists of 491 nodes and 3390 edges. The network node consists of 117 compounds and 264 targets (Fig. 1E). It can be known from the C-T network that different nodes play different roles in the network. For example, some target nodes have large degree value in the network, indicating that there are more compounds involved in regulating the target which means that the target is more strongly regulated by Shenlian Capsule. For instance, the average degree value of all regulated targets is 10.50, but the degree values of the top five regulated target nodes are: AR: 130, ESR1: 124, PTGS2: 107, PIM1: 103, CDK2: 99.

Shenlian Capsule C-D-T network analysis

The network analysis plug-in of Cytoscape software were used to intersect the C-T network with the D-T network (Fig. 2B) to obtain a total of 47 targets (Fig. 2C). The C-D-T network was constructed by combining intersection targets with compounds (Fig. 2D). The C-D-T network consists of 141 nodes and 541 edges. The nodes are composed of 47 NSCLC related targets and 94 potential compound nodes. By analyzing the C-T-D network, 11 disease-related targets were be found with degree value greater than 10, which PTGS2, CDK2, JUN, TNF, BCL2, IL6, BAX, AKT1, PPARG, HMOX1 and MMP1 target proteins were sequentially 107, 97, 19, 17, 15, 14, 13, 12, 12, 11, 11 regulatory numbers of potential active compounds. In addition, from the C-D-T network, potential active compounds in Shenlian Capsule can regulate multiple protein targets. For example, Quercetin can regulate 39 disease-related targets, luteolin regulates 24 disease-related targets and kaempferol regulate 15 Disease-related targets. Through analysis C-D-T network find that Shenlian Capsule can treat NSCLC through multi-component and multi-target.

Analysis of target protein enrichment based on C-D-T network

By enriching the biological process of target proteins in the C-D-T network, Shenlian Capsule can regulate angiogenesis, nitric oxide synthesis, cell proliferation, cell apoptosis, and cell response to DNA damage stimuli, ERK1 and ERK2 Positive cascade regulation, protein phosphorylation, ERBB2 signaling pathway and other processes to fight NSCLC. Through molecular function analysis of Go enrichment analysis, it was found that the target nodes in the C-D-T network are mainly related to enzyme binding, transcription factor binding, protein phosphatase binding, receptor signaling protein tyrosine kinase activity and other functions. The cellular component analysis of the Go enrichment analysis found that the target nodes in the C-D-T network mainly acted on the nucleus and cytosol of the cells (Fig. 2E).

According to pathway enrichment analysis of 47 disease-related proteins, it can be seen that Shenlian Capsule are enriched in multiple pathways. Such as Pathways in cancer, T cell receptor signaling pathway, B cell receptor signaling pathway, TNF signaling pathway, NF-kappa B signaling pathway, PPAR signaling pathway, p53 signaling pathway, Wnt signaling pathway, VEGF signaling pathway, Thyroid hormone signaling pathway, Sphingolipid signaling pathway, MAPK signaling pathway, Ras signaling pathway, PI3K-Akt signaling pathway and other pathways (Fig. 2F).

Cluster Analysis of Core Targets for PPI Network

Targets in the C-D network were placed in Bisogenet to construct a PPI network for the target of Shenlian Capsule. The network consists of 8442 target nodes and 186422 interactions (Fig. 3A-a). The NSCLC

related targets PPI network constructed as described above, the target network and a total of 6,554 targets and 160,234 edges (Fig. 3A-b).

The drug PPI network is further intersected with the disease PPI network to obtain the C-D-T PPI network. The network consists of 5,574 nodes and 143,784 interactions (Fig. 3B-a). In order to obtain the core targets of drug-acting diseases, the targets were first screened with a median value greater than 2 times the degree, and a network of 1367 nodes and 60209 edges was obtained (Fig. 3B-b). The core targets network was further screened by topology method, and the core targets of drug affected diseases were screened by used degree > 136, BC > 555.53, EC > 0.0160, LAC > 15.205 and NC > 127.70. The network consists of only 102 nodes and 2442 edges (Fig. 3B-c).

Table 1
Clustering Analysis Results of Core Targets in C-D-T PPI Network

Cluster	Score	Nodes	Edges	Node name
1	19.91	22	209	RPS7, RPS8, NPM1, HNRNPA1, RPS6, RPS3, RPL11, RPS2, RPS4X, RPL13, RPS3A, RPLP0, RPSA, RPL10, RPL5, HNRNPU, RPS14, NTRK1, RPS16, VCAM1, RPL23, UBL4A
2	16.91	36	296	RACK1, CCDC8, TARDBP, CUL2, CUL3, CUL1, MCM2, PABPC1, NOP56, U2AF2, DHX9, CUL5,PAN2, HIST1H3H, FN1, CAND1, HIST1H3J, HIST1H3F, ILF3, ILF2, HIST1H3B, HIST1H3E, HIST1H3C, HIST1H3G, HIST1H3I, HIST1H3D, HIST1H3A, STAU1, NCL, RPL6, YWHAG, HDAC1, FUS, EP300, HSP90AA1, PARP1
3	12.62	30	183	ITGA4, HNRNPD, YWHAQ, CDK2, SNW1, SUZ12, TUBG1, ESR1, NEDD8, VCP, CUL4B,EEF1A1, EGFR, UBC, CUL7, CDC5L, SYNCRIP, OBSL1, SMARCA4, YWHAZ, HDAC2, RPA1, HUWE1, XPO1, HSPA8, HNRNPM, EIF4A3, MCM5, HSP90AB1, CREBBP,
4	3.56	10	16	HNRNPK, SRRM2, GRB2, RNF2, YWHAЕ, TUBB, RPS27A, DDX5, HDAC5, HSPA5

Four clusters were obtained by clustering the core targets of the C-D-T PPI network by used MCODE plug-in. Cluster1 is consist of 22 nodes and 229 edges, cluster2 is consist of 36 nodes and 296 edges, cluster3 is consist of 30 nodes and 183 edges, cluster4 is consist of 10 nodes and 16 edges (Fig. 3C and Table 1).

Cluster target GO enrichment and pathway analysis

To illustrate the biological functions of the 102 major targets, through classified the core PPI network of drug diseases into 4 clusters (Fig. 4A), performed GO (Fig. 4B) and KEGG pathway (Fig. 4C) enrichment analysis on each cluster.

Based on these GO terms enrichment analysis data found that (a) gene expression, RNA binding, rRNA binding, protein kinase binding, skinase binding;(b) negative regulation of cell proliferation, protein monoubiquitination, ubiquitin protein ligase activity, receptor tyrosine kinase binding; (c) DNA repair, DNA replication;(d) positive regulation of transcription from RNA polymerase II promoter, negative regulation of epidermal growth factor receptor signaling pathway, ERBB2 signaling pathway, negative regulation of

apoptotic process, ubiquitin protein ligase binding, cadherin binding involved in cell-cell adhesion. Enrichment analysis by Go suggesting that Shenlian Capsule treat NSCLC from multiple angles.

By analyzing the KEGG pathway data of each pathway, we consider that the key signaling pathways of Ribosome, Cell cycle, Viral carcinogenesis, PI3K-Akt signaling pathway, Notch signaling pathway and FoxO signaling pathway, the top 6 KEGG pathways in four clusters, might be the most important pharmacological mechanism of Shenlian Capsule for NSCLC.

Effect of Shenlian Capsule on the role of LUSC survival-related gene

LUSC Differential gene

Through differential gene analysis, we obtained a heatmap of the differential genes of LUSC disease (Fig. 5A-a) and a volcano map of the expression levels of differential genes (Fig. 5A-b). A total of 4,664 differential genes which 2836 genes were up-regulated and 1,828 genes were down-regulated.

Survival related genes of LUSC treated by Shenlian Capsule

A survival analysis was performed for each of the above-mentioned differential genes and a total of 256 differences were acquired. The 256 differences were intersected with the C-D network to obtain CTSD and PLAU survival-related genes (Fig. 5A-c). CTSD is regulated by MOL000098 active compounds and PLAU is regulated by two active compounds.

Differential expression of normal tissue and cancer tissue of LUSC survival related genes treated by Shenlian Capsule

CTSD expression in normal tissue was 594.57 ± 235.35 , cancer tissue expression was 293.66 ± 165.48 , cancer tissue expression was significantly lower than normal tissue, $p < 0.01$ (Fig. 5B-a). The expression of PLAU in normal tissue was 12.17 ± 6.69 and cancer tissue was 87.05 ± 92.70 , the expression of PLAU cancer tissue was higher than normal tissue, $p < 0.01$ (Fig. 5B-c).

Survival curve of LUSC survival related genes treated by Shenlian Capsule

The CTSD gene survival curve was significantly different in the low expression group and the higher expression group, $p < 0.01$ (Fig. 5B-b). The survival time of the low expression group of the PLAU survival curve is better than that of the high expression group, $p < 0.01$ (Fig. 5B-d).

Molecular docking verify the effect of active compounds on survival-related targets

After docking the MOL00098 active compound with the CTSD protein, the selected conformational energy is -7.1Kcal/mol. The active compound has five hydrogen bonds to interact with the CTSD protein. From the two-dimensional structure after docking, it can be seen that the active compound has two hydrogen bonds with the Tyr residue at position 205 and the distance are 2.77 Å. In addition, the active compound has hydrogen bonding interactions with Ser residue at position 80, Ile residue at position 142 and Gly residue at position 233. The hydrogen bond distances are 2.90 Å, 2.81 Å and 2.82 Å (Fig. 5C-a). Figure 5C-b shows the three-dimensional structure diagram of the active compound MOL00098 after docking with CTSD, and the two-dimensional structure diagram after docking produces good mutual confirmation.

After docking the MOL000449 active compound with PLAU protein, the selected conformational energy is

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js is to interact with the PLAU protein. According to

the two-dimensional structure after docking, it can be seen that the active compound has 2 hydrogen bonds with the Arg residue at position 239, the distances are 2.97 Å and 3.05 Å, respectively. In addition, the active compound has a weak hydrogen bond interaction with the Cys residue at position 13 with a hydrogen bond distance of 3.27 Å (Fig. 5C-c). Figure 5C-d is a three-dimensional structure diagram of the docking of active compound MOL00449 and PLAU, showing the most stable conformation of the docking of active compound with PLAU protein.

Effect of Shenlian Capsule on the role of LUAD survival-related targets

LUAD Differential gene

Through differential gene analysis obtained a heat map (Fig. 6A-a) of the differential genes of LUAD and a volcano map (Fig. 6A-b) of the expression levels of differential genes. A total of 6776 differential genes which 5176 genes were up-regulated and 1600 genes were down-regulated.

Survival related genes of LUAD treated by Shenlian Capsule

A survival analysis was performed for each of the above-mentioned differential genes, and a total of 490 difference genes were obtained. Intersecting 490 difference genes with the C-D network yielded eight survival-related genes for ADRB2, BIRC5, CCNA2, CCNB1, CDK1, CHEK1, F2 and MMP3 (Figs. 6A-c). The above eight survival-related genes are regulated by 46, 74, 3, 3, 58, 72 and 1 active compounds, respectively.

Differential expression of normal tissue and cancer tissue of LUAD survival related genes treated by Shenlian Capsule

The ADRB2 expression in normal tissue was 12.80 ± 5.09 , cancer tissue expression was 1.71 ± 1.59 and cancer tissue expression was significantly lower than normal tissue, $p < 0.01$ (Fig. 6C-a). The expression of BIRC5 in normal tissue was 0.81 ± 0.60 and cancer tissue was 11.46 ± 10.41 , cancer tissue is higher than normal tissues, $p < 0.01$ (Fig. 6C-b). The expression of CCNA2 in normal tissue was 0.97 ± 0.56 and cancer tissue is 7.97 ± 7.23 , the expression of cancer tissue is higher than normal tissues, $p < 0.01$ (Fig. 6C-c). The expression of CCNB1 in normal tissue was 2.01 ± 0.91 and cancer tissue is 20.97 ± 14.07 , the expression of cancer tissue is higher than normal tissues, $p < 0.01$ (Fig. 6C-d). The expression of CDK1 in normal tissue was 1.56 ± 0.57 and cancer tissue was 9.51 ± 8.46 , the expression of cancer tissue is higher than normal tissues, $p < 0.01$ (Fig. 6C-e). The expression of CHEK1 in normal tissue was 0.54 ± 0.13 and cancer tissue was 2.82 ± 2.30 , the expression of cancer tissue is higher than normal tissues, $p < 0.01$ (Fig. 6C-f). The expression of F2 in normal tissue was 0.0045 ± 0.0074 and cancer tissue was 0.80 ± 7.13 , the expression of cancer tissue is higher than normal tissues, $p < 0.01$ (Fig. 6C-g). The expression of MMP3 in normal tissue was 0.074 ± 0.11 and cancer tissue was 1.26 ± 2.87 , the expression of cancer tissue is higher than normal tissues, $p < 0.01$ (Fig. 6C-h).

Survival curve of LUAD survival related genes treated by Shenlian Capsule

The ADRB2 survival curve was significantly different in the low expression group and the higher expression group, $p < 0.01$ (Fig. 6D-a). The greater the number of ADRB2 gene expressions, the longer the patient's survival time. The survival time of the low expression group of the BIRC5, CCNA2, CCNB1, CDK1, the high expression group, $p < 0.01$ (Fig. 6D(b-h)).

The lower the amount of tumor tissue expression of the above seven genes, the longer the patient's survival time.

Molecular docking used to verify the effect of active compounds on survival-related genes

Forty-six active compounds were docked with the ADRB2 protein, the lower energy conformation was selected for display. The result was the conformation with the lower energy after the MOL07111 active compound was docked with the ADRB2 protein, the selected conformation energy is -7.8Kcal/mol. The active compound has 3 hydrogen bonds to interact with the protein. From the two-dimensional and three-dimensional structure after docking, it can be seen that the active compound has two hydrogen bonds at the amino terminal of the Tyr residue at the O1 position and the amino terminal of the Gly residue at the 25 position. The two hydrogen bonding distances are 3.07 Å and 3.24 Å. In addition, the O2 terminal of the active compound has a hydrogen bonding interaction with the amino terminal of the Arg residue at position 80 and distance is 3.00 Å (Fig. 6E-a, Fig. 6E-b).

Three active compounds were docked with the BIRC5 protein, the conformation with the lower energy for docking MOL00098 active compounds with the BIRC5 protein was been selected. Its conformational energy was - 7.1Kcal/mol. The active compound has 3 hydrogen bonds to interact with the BIRC5 protein. From the two-dimensional and three-dimensional structures after docking, it can be seen that the active compound has two hydrogen bonds at the amino terminal of the 16 position Asp residue at the O4 position and the amino terminal of the Lys residue at the 15 position. The distances between the two hydrogen bonds are 2.71 Å and 3.21 Å. In addition, the O6 terminal of the active compound has a hydrogen bonding interaction with the amino terminal of the Val residue at position 89 and distance is 2.74 Å (Fig. 6E-b, Fig. 6F-b).

Seventy-eight active compounds were docked with the CCNA2 protein, the conformation with the lower energy for docking MOL07150 active compounds with the CCNA2 protein was been selected. Its conformational energy was - 10.0 Kcal/mol. The active compound has 4 hydrogen bonds to interact with the protein. From the two-dimensional and three-dimensional structures after docking, the O4 terminal of the MOL07150 compound has two hydrogen bonds with the hydroxyl position of the Gln residue at position 81 and the hydroxyl position of the Tyr residue at position 113. The hydrogen bond distances are 3.15 Å and 3.28 Å, respectively. The O5 terminal of the MOL07150 compound has two hydrogen bonds with the hydroxyl position of the Lys residue at position 78 and the hydroxyl position of the Tyr residue at position 113, the hydrogen bond distances are 2.77 Å and 2.99 Å (Fig. 6E-c, Fig. 6F-c).

Three active compounds were docked with the CCNB1 protein, the conformation with the lower energy for docking MOL002714 active compounds with the CCNB1 protein was been selected. Its conformational energy was - 7.0Kcal/mol. From the two-dimensional and three-dimensional structures after docking, the O2 terminal of the MOL002714 compound has a hydrogen bond interaction with the amino terminal of the Val residue at position 156 and distance is 3.12 Å (Fig. 6E-d, Fig. 6F-d).

Three active compounds were docked with the CDK1 protein, and the conformation with the lower energy for docking MOL002714 active compounds with the CDK1 protein was been selected. Its conformational

energy was -9.6 Kcal/mol. The active compound has 4 hydrogen bonds to interact with the CDK1 protein. It can be seen from the two-dimensional and three-dimensional structures after docking that the active compound has two hydrogen bonds at the O4 position with the amino terminal and the hydroxyl terminal of the 12 position Glu residue. The distance between the two hydrogen bonds are 2.88 Å and 3.21 Å. In addition, the O5 terminal of the active compound produces a hydrogen bond interaction with the hydroxyl end of the Ile residue at position 10 and distance is 3.18 Å. The O7 terminal of the active compound produces a hydrogen bond interaction with the amino terminus of the Asp residue at position 146 and distance is 2.93 Å (Fig. 6E-e, Fig. 6F-e).

Fifty-eight active compounds were docked with the CHEK1 protein, the conformation with the lower energy for docking MOL002651 active compounds with the CHEK1 protein was been selected. Its conformational energy was -9.7 Kcal/mol. It can be seen from the two-dimensional and three-dimensional structures after docking that the oxygen at the 3 position of the active compound has a weak hydrogen bond interaction with the amino terminus of the Cys residue at the 87 position and distance is 3.25 Å. In addition, the active compound has significant hydrophobic interactions with the surrounding 11 amino acids (Fig. 6E-f, Fig. 6F-f).

Seventy-two active compounds were docked with the F2 protein, the conformation with the lower energy for docking MOL007151 active compounds with the F2 protein was been selected. Its conformational energy was -7.1 Kcal/mol. It can be seen from the two-dimensional and three-dimensional structures after docking that O4 and O5 terminal of the active compound with Glu residue at position 362 of the F2 protein have two hydrogen bonding interactions, the distance are 2.71 Å and 2.91 Å. In addition, the active compound has significant hydrophobic interactions with the surrounding 7 amino acid residues.

After docking the active compound MOL00098 with MMP3, the lower energy conformation was been selected, the energy after docking was -7.3 Kcal/mol. It can be seen from the two-dimensional and three-dimensional structures after docking that O4 terminal of the active compound has hydrogen bonding interactions with the Tyr residue at position 220 and the Leu residue at position 222 of the MMP3 protein, the distances are 2.70 Å and 3.23 Å. In addition, the active compound has significant hydrophobic interactions with the surrounding 6 amino acids.

Discussion

From system pharmacological research found that Shenlian Capsule acted on multiple targets for the treatment of NSCLC through multiple active ingredients. From the C-D-T network, Shenlian Capsule were found to act on 47 NSCLC associated targets. The enrichment analysis of the 47 targets revealed that Shenlian Capsule involved multiple pathways in the treatment of NSCLC. These cellular pathways are related to anti-inflammatory, immune, vascular proliferation, apoptosis, Hypoxic adaptation, cell proliferation and other processes. In the C-D-T network, the enrichment of HIF-1 signaling pathway, ErbB signaling pathway, PI3K-Akt signaling pathway and FoxO signaling pathway are relatively high.

When solid tumors grow rapidly, the microenvironment of tumor cells is likely to cause hypoxia, tumor cells need to adapt the hypoxic environment to continue growing. HIF-1 signaling pathway is a key pathway for tumor cells to adapt the hypoxic microenvironment to continue to grow. At the same time, the adaptation of tumor cells to the hypoxic environment is also an important reason for the resistance of NSCLC to chemotherapy and radiotherapy. Patients with high HIF-1 α expression levels have significantly shorter survival times and farther metastases. The expression of HIF-1 α in tumor tissues of non-small cell patients is higher than that of normal tissues[33–36]. This study found that the active compounds of Shenlian Capsule act on 12 proteins related to HIF-1 signaling pathway.

EGFR (Epidermal Growth Factor Receptor) is a receptor for epithelial growth factor (EGF) cell proliferation and signaling. EGFR belongs to a family of ErbB receptors. EGFR is also known as HER1, ErbB1, which is often mutated or overexpressed in tumor tissues. EGFR tyrosine kinase inhibitors can inhibit EGFR expression. A variety of EGFR tyrosine kinase inhibitors have been widely used in clinical or ongoing clinical trials, such as gefitinib, erlotinib, vandetanib (ZD-6474) and monoclonal antibodies[37]. EGFR-mediated tumor progression, such as tumor cell proliferation, differentiation, migration and invasion[38]. This study found that the active compounds of Shenlian Capsule act on 11 proteins related to ErbB signaling pathway.

The phosphatidylinositol 3 kinase (PI3K) / AKT (protein kinase B) signaling pathway can be activated by various types of cell stimuli and toxins, it also can participate in the regulation of many basic cellular processes, including cell growth, transcription, translation, proliferation, cell movement and glycogen metabolism[39]. The PI3K / AKT pathway has been altered in a variety of cancers, including NSCLC. Mutations and overexpression are more common in LUSC than LUAD [40]. In addition, abnormal activation of the PI3K/AKT pathway is one of the important mechanisms for adenocarcinoma patients to develop resistance to EGFR-TK inhibitors[41]. The AKT inhibitors currently in clinical use or participating in clinical trials include rapamycin and raparos [40]. This study found that the active compounds of Shenlian Capsule act on 17 proteins related to PI3K/ AKT signaling pathway.

FoxO signaling pathway abnormalities have a wide range of effects on cell biological processes, including cell cycle regulation, apoptosis, DNA damage repair, involvement as an oxidative stress mediator, angiogenesis and tumorigenesis[42]. The expression of FoxO protein in lung cancer tissue was less than that in normal tissue [43]. FoxO1 negatively regulates the growth of lung cancer tumor cells in mice[44]. Increased FoxO3 expression enhances radiation resistance in NSCLC cells[45]. This study found that the active ingredients of Shenlian Capsule act on 17 proteins related to the FoxO signaling pathway.

Through clustering and enrichment analysis of core targets, 6 pathways including Viral carcinogenesis and Notch signaling pathway were obtained. Many studies have shown that viruses (mainly with human papillomavirus (HPV)) are important factors in NSCLC[46]. Studies have now shown that the virus plays an important role in the development of lung cancer, with exogenous retroviruses found in 85% of squamous cell carcinomas and 47% of adenocarcinomas, approximately 30% of squamous cell carcinomas in samples contain high-risk HPV [47]. Oncoviruses play a role in NSCLC by regulating host

genes in tumor cells during NSCLC differentiation and development[48]. Our research shows that 4,6 proteins in cluster3 and cluster4 of the core targets are related to Viral carcinogenesis, respectively.

Notch signaling pathway members have abnormal expression in NSCLC. Notch signaling pathway plays an important role in regulating the proliferation, differentiation and apoptosis of NSCLC. In addition, inhibiting Notch signaling pathway can inhibit lung tumor cell apoptosis. Overexpression of Notch1 and Notch3 protein molecules and increased lymph node metastasis in NSCLC patients lead to shortened patient survival time[49]. Our research shows that 3 proteins in cluster3 of the core targets are related to Notch signaling pathway.

We finally used bioinformatics methods to study LUSC and LUAD survival-related genes in NSCLC, by intersected the survival-related genes with the C-T network. It was obtained that Shenlian Capsule have LUSC 2 survival-related genes and LUAD have 8 survival-related genes (Fig. 7). The expression levels of the 10 survival-related genes in the cancer group were significantly different from normal tissues. In order to verify the reliability of the WES method and the prediction of SysDT models, we performed molecular docking of active compounds acting on 10 survival-related genes, used two-dimensional and three-dimensional diagrams to display the docking results, demonstrating that Shenlian Capsule acts on 10 survival-related genes.

Most of these 10 survival-related genes have been reported in previous research centers to be related to the mechanism of NSCLC disease. PLAU protein is a component of urokinase-type plasminogen activator. Studies have shown that the expression of this complex in lung cancer tissue is higher than that in normal tissues[50]. Our research also proved that PLAU expression in lung cancer tissue was significantly increased. The complex enzyme can promote the migration and infiltration of tumor cells by destroying the cell matrix[51]. CTSD protein is a kind of cathepsin. CTSD can directly catalyze the degradation of extracellular matrix and extrabasal membrane, participate in the proteolytic cascade, jointly promote tumor growth, invasion and metastasis[52]. Our research shows that CTSD expression is inversely related to patient survival time. It is known through molecular docking that the related active compounds of Shenlian Capsule can obviously act on CTSD and PLAU proteins.

Previous studies have shown that ADRB2 expression in lung tumor cells is significantly lower than normal tissues[53], which is consistent with our study of the differences between normal and tumor tissues of this protein. ADRB2 receptor inhibitors are associated with lung tumor tissue proliferation, angiogenesis, migration and invasion, immune invasion and survival[54]. Some studies have shown that the expression of BIRC5 in lung cancer tissue is significantly higher than that of tissues adjacent to lung cancer[55], which is consistent with our research. BIRC5 has the functions of inhibiting apoptosis, promoting cell transformation, participating in cell mitosis, angiogenesis and drug resistance[56].

Lung cancer tissue have faster division of cancer cells and a shorter cell cycle. CDKs protein activation is a key driver of cell cycle transition in lung cancer cells[57]. CCNB1 overexpression in NSCLC is closely related to tumor type, tumor differentiation, vascular infiltration and positive[58]. CCNA2 overexpression

Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js

es the migration and invasiveness of NSCLC

cells[59]. Studies have shown that CCNA2, CCNB1, CDKs and other regulatory cyclin expression are higher than adjacent tissues, and CCNA2, CCNB1, CDKs protein expression levels are consistent with our research[60]. In addition, this work found that Shenlian Capsule have multiple compounds that regulate the three cell cycle related proteins. CHEK1 encodes a serine / threonine kinase, which is a key component of cells' response to DNA damage. CHEK1 participates in the repair of cell cycle arrest, programmed cell death and other processes by regulating cell cycle checkpoints[61]. Inhibition of CHEK1 can increase the lethality of chemotherapeutic drugs to tumor cells and reduce the toxicity of chemotherapy to normal tissues[62].

Coagulation factor 2(F2) is an important component of thrombin. Thrombin regulates multiple processes of cancer spread such as: invasion, migration, access to blood vessels, survival in the lymphatic system and promoting tumor growth[63]. By using a heparin derivative and heparanase inhibitor prothrombin inhibitory mechanism, tumor growth and neovascularization can be inhibited[64]. In our research center, the higher expression of the F2 low expression group was significantly improved. Matrix metalloproteinase 3 (MMP3) can degrade, reshape extracellular matrix and fully activate multiple matrix metalloproteinase precursors. Its increased expression can promote the development of a variety of diseases, especially metastatic diseases such as tumors and inflammatory infiltration. Increased expression of MMP3 caused by obesity can promote lung cancer metastasis by activating MMP9[65]. Our research shows that the expression of MMP3 in cancer tissue is significantly higher than in cancerous fat tissues and the survival curve of the low expression group is better than that of the high expression group. The molecular docking shows that the active compound of Shenlian Capsule has lower energy and is more stable after docking with the above 8 proteins related to the survival of LUAD.

Conclusion

This study uses a system pharmacological approach to study the Shenlian Capsule in the treatment of NSCLC. Its main research methods include network pharmacology method, topology method, bioinformatics method and molecular docking technology. Through the network pharmacological method found that Shenlian Capsule can act on 47 target proteins related to NSCLC. According to the enrichment analysis, it is known that the 47 targets are involved in anti-inflammatory, immunity, angiogenesis, hypoxia stress, apoptosis, antioxidant, cell proliferation and other processes. This work screened the core targets by topology method found that Shenlian Capsule in treating NSCLC may be related to the carcinogenic pathway of the virus. Using bioinformatics methods, Shenlian Capsule were found to act on LUSC survival-related 2 proteins and LUAD survival-related 8 proteins. The expression of these 10 proteins was significantly different in tumor tissues and adjacent tissues and there were also significant differences in the survival curve low expression group and high expression group. Through the molecular docking technology, we can better show the two-dimensional and three-dimensional structure diagrams of the lower energy conformation of 10 survival-related genes after docking with the active compound of Shenlian Capsule.

Abbreviations

ADME

Absorption, Distribution, Metabolism and Excretion; BBD:BaiBianDou, BC:betweenness centrality; BGZ:BuGuZhi; BIND:Biomolecular Interaction Network Database; BioGRID:Biological General Repository for Interaction Datasets; BP:biological process; BZL:BanZhiLian; CC:closeness centrality; C-D-T:compounddisease-target; CFDA:China Food and Drug Administration; C-T:compounds target; DC:degree centrality; DIP:Database of Interacting Proteins; DL:drug-like; DS:DanShen; D-T:disease-target; EC:eigenvector centrality; EZ:EZhu; FJ:FangJi; HL:half-life; HPRD:Human Protein Reference Database; IntAct:InterAct Molecular Interaction Database; KM:Kaplan-Meier; KS:KuShen;KXR:KuXingRen;LAC:largest average local connectivity; LUAD:Non-Small Cell Lung Adenocarcinoma; LUSC:Non-Small Cell Squamous Cell Lung Carcinoma; MF:molecular functions; MINT:Molecular INTERaction Database; NC:network centrality; NSCLC:NonSmall Cell Lung Cancer; OB:oral bioavailability; OMIM:Online Mendelian Inheritance in Man; PPI:protein-protein interaction; SDG:ShanDouGen; SL:SanLeng; TCM:Traditional Chinese medicine; TTD:Therapeutic Target Database; WM:WuMei.

Declarations

Acknowledgements

This study was supported by grants from the Natural Science Foundation of Anhui Province (1908085QH317).

Author Contributions

WF and AG made the same contribution in this work. Investigation, WF; methodology, WF,XW,YC,BY; project administration, WH; resources, AG; software, WF; supervision, WH; writing (original draft), AG; writing (review & editing), RZ and YC.. The author(s) read and approved the final manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Author details

¹Department of Clinical Pharmacology, The Second Hospital of Anhui Medical University, Hefei,230601, China.²Department of Endocrinology, The Second Hospital of Anhui Medical University, Hefei,230601, China.

References

1. Barta JA, Powell CA, Wisnivesky JP. Global Epidemiology of Lung Cancer. ANN GLOB HEALTH.

2. Cao M, Chen W. Epidemiology of lung cancer in China. *THORAC CANCER*. 2019;10:3–7.
3. Lu T, Yang X, Huang Y, Zhao M, Li M, Ma K, et al. Trends in the incidence, treatment, and survival of patients with lung cancer in the last four decades. *CANCER MANAG RES*. 2019;11:943–53.
4. Torre LA, Siegel RL, Jemal A. Lung Cancer Statistics. *ADV EXP MED BIOL*. 2016;893:1–19.
5. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. *NATURE*. 2018;553:446–54.
6. ZOU Zhong-li LXLC. Clinical study on Shenlian Capsules combined with tegafur, gimeracil and oteracil potassium in treatment of advanced gastric cancer. *Drugs Clinic*. 2019;05:1520–4.
7. GE L. Clinical observation of non-small cell lung cancer treated by erlotinib and senlian capsule. *Strait Pharmaceutical Journal*. 2010;22: 96–7.
8. Hu W, Fu W, Wei X, Yang Y, Lu C, Liu Z. A Network Pharmacology Study on the Active Ingredients and Potential Targets of *Tripterygium wilfordii* Hook for Treatment of Rheumatoid Arthritis. *Evid Based Complement Alternat Med*. 2019;2019:5276865.
9. Zhang R, Zhu X, Bai H, Ning K. Network Pharmacology Databases for Traditional Chinese Medicine: Review and Assessment. *FRONT PHARMACOL*. 2019;10:123.
10. Zhang W, Bai Y, Wang Y, Xiao W. Polypharmacology in Drug Discovery: A Review from Systems Pharmacology Perspective. *Curr Pharm Des*. 2016;22:3171–81.
11. Chen CY. TCM Database@Taiwan: the world's largest traditional Chinese medicine database for drug screening in silico. *PLOS ONE*. 2011;6:e15939.
12. Ru J, Li P, Wang J, Zhou W, Li B, Huang C, et al. TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. *J Cheminform*. 2014;6:13.
13. Tao W, Li B, Gao S, Bai Y, Shar PA, Zhang W, et al. CancerHSP: anticancer herbs database of systems pharmacology. *Sci Rep*. 2015;5:11481.
14. Jiang Y, Zhong M, Long F, Yang R, Zhang Y, Liu T. Network Pharmacology-Based Prediction of Active Ingredients and Mechanisms of *Lamiophlomis rotata* (Benth.) Kudo Against Rheumatoid Arthritis. *FRONT PHARMACOL*. 2019;10:1435.
15. Xu X, Zhang W, Huang C, Li Y, Yu H, Wang Y, et al. A novel chemometric method for the prediction of human oral bioavailability. *INT J MOL SCI*. 2012;13:6964–82.
16. Tian S, Wang J, Li Y, Xu X, Hou T. Drug-likeness analysis of traditional Chinese medicines: prediction of drug-likeness using machine learning approaches. *Mol Pharm*. 2012;9:2875–86.
17. Tao W, Xu X, Wang X, Li B, Wang Y, Li Y, et al. Network pharmacology-based prediction of the active ingredients and potential targets of Chinese herbal *Radix Curcumae* formula for application to cardiovascular disease. *J ETHNOPHARMACOL*. 2013;145:1–10.
18. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *GENOME RES*. 2003;13:2498–504.

19. Martin A, Ochagavia ME, Rabasa LC, Miranda J, Fernandez-de-Cossio J, Bringas R. BisoGenet: a new tool for gene network building, visualization and analysis. BMC BIOINFORMATICS. 2010;11:91.
20. Tang Y, Li M, Wang J, Pan Y, Wu FX. CytoNCA: a cytoscape plugin for centrality analysis and evaluation of protein interaction networks. BIOSYSTEMS. 2015;127:67–72.
21. Song W, Ni S, Fu Y, Wang Y. Uncovering the mechanism of Moxing Ganshi Decoction on asthma from a systematic perspective: A network pharmacology study. Sci Rep. 2018;8:17362.
22. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. BMC BIOINFORMATICS. 2003;4:2.
23. Wang J, Zhong J, Chen G, Li M, Wu FX, Pan Y. ClusterViz. A Cytoscape APP for Cluster Analysis of Biological Network. IEEE/ACM Trans Comput Biol Bioinform. 2015;12:815–22.
24. Huang DW, Sherman BT, Lempicki RA. Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. NUCLEIC ACIDS RES. 2009;37:1–13.
25. Namani A, Zheng Z, Wang XJ, Tang X. Systematic Identification of Multi Omics-based Biomarkers in KEAP1 Mutated TCGA Lung Adenocarcinoma. J CANCER. 2019;10:6813–21.
26. Cheng X, Yin H, Fu J, Chen C, An J, Guan J, et al. Aggregate analysis based on TCGA: TTN missense mutation correlates with favorable prognosis in lung squamous cell carcinoma. J Cancer Res Clin Oncol. 2019;145:1027–35.
27. Cirenajwis H, Lauss M, Planck M, Vallon-Christersson J, Staaf J. Performance of gene expression-based single sample predictors for assessment of clinicopathological subgroups and molecular subtypes in cancers: a case comparison study in non-small cell lung cancer. BRIEF BIOINFORM. 2019.
28. beeswarm. The bee swarm plot, an alternative to stripchart. 0.1. 5. 2012.
29. Trott O, Olson AJ. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J COMPUT CHEM. 2010;31:455–61.
30. Gao YD, Huang JF. [An extension strategy of Discovery Studio 2.0 for non-bonded interaction energy automatic calculation at the residue level]. Dongwuxue Yanjiu. 2011;32:262–6.
31. Laskowski RA, Swindells MB. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. J CHEM INF MODEL. 2011;51:2778–86.
32. Janson G, Zhang C, Prado MG, Paiardini A. PyMod 2.0: improvements in protein sequence-structure analysis and homology modeling within PyMOL. BIOINFORMATICS. 2017;33:444–6.
33. Wang G, Xiao L, Wang F, Yang J, Yang L, Zhao Y, et al. Hypoxia inducible factor-1alpha/B-cell lymphoma 2 signaling impacts radiosensitivity of H1299 non-small cell lung cancer cells in a normoxic environment. Radiat Environ Biophys. 2019;58:439–48.
34. Wang Q, Hu DF, Rui Y, Jiang AB, Liu ZL, Huang LN. Prognosis value of HIF-1alpha expression in patients with non-small cell lung cancer. GENE. 2014;541:69–74.
35. Kim SJ, Rabbani ZN, Dewhirst MW, Vujaskovic Z, Vollmer RT, Schreiber EG, et al. Expression of HIF-1alpha, CA IX, VEGF and MMP-9 in surgically resected non-small cell lung cancer. LUNG CANCER.

2005;49:325–35.

36. Sobhanifar S, Aquino-Parsons C, Stanbridge EJ, Olive P. Reduced expression of hypoxia-inducible factor-1alpha in perinecrotic regions of solid tumors. *CANCER RES.* 2005;65:7259–66.
37. Cascone T, Martinelli E, Morelli MP, Morgillo F, Troiani T, Ciardiello F. Epidermal growth factor receptor inhibitors in non-small-cell lung cancer. *Expert Opin Drug Discov.* 2007;2:335–48.
38. Weng MS, Chang JH, Hung WY, Yang YC, Chien MH. The interplay of reactive oxygen species and the epidermal growth factor receptor in tumor progression and drug resistance. *J Exp Clin Cancer Res.* 2018;37:61.
39. Martini M, De Santis MC, Braccini L, Gulluni F, Hirsch E. PI3K/AKT signaling pathway and cancer: an updated review. *ANN MED.* 2014;46:372–83.
40. Fumarola C, Bonelli MA, Petronini PG, Alfieri RR. Targeting PI3K/AKT/mTOR pathway in non small cell lung cancer. *BIOCHEM PHARMACOL.* 2014;90:197–207.
41. Singh M, Jadhav HR. Targeting non-small cell lung cancer with small-molecule EGFR tyrosine kinase inhibitors. *DRUG DISCOV TODAY.* 2018;23:745–53.
42. Farhan M, Wang H, Gaur U, Little PJ, Xu J, Zheng W. FOXO Signaling Pathways as Therapeutic Targets in Cancer. *INT J BIOL SCI.* 2017;13:815–27.
43. Maekawa T, Maniwa Y, Doi T, Nishio W, Yoshimura M, Ohbayashi C, et al. Expression and localization of FOXO1 in non-small cell lung cancer. *ONCOL REP.* 2009;22:57–64.
44. Wang C, Liu E, Li W, Cui J, Li T. MiR-3188 Inhibits Non-small Cell Lung Cancer Cell Proliferation Through FOXO1-Mediated mTOR-p-PI3K/AKT-c-JUN Signaling Pathway. *FRONT PHARMACOL.* 2018;9:1362.
45. Gao Z, Liu R, Ye N, Liu C, Li X, Guo X, et al. FOXO1 Inhibits Tumor Cell Migration via Regulating Cell Surface Morphology in Non-Small Cell Lung Cancer Cells. *CELL PHYSIOL BIOCHEM.* 2018;48:138–48.
46. Klein F, Amin KW, Petersen I. Incidence of human papilloma virus in lung cancer. *LUNG CANCER.* 2009;65:13–8.
47. Robinson LA, Jaing CJ, Pierce CC, Magliocco A, Xiong Y, Magliocco G, et al. Molecular evidence of viral DNA in non-small cell lung cancer and non-neoplastic lung. *Br J Cancer.* 2016;115:497–504.
48. Kim Y, Pierce CM, Robinson LA. Impact of viral presence in tumor on gene expression in non-small cell lung cancer. *BMC CANCER.* 2018;18:843.
49. Zou B, Zhou XL, Lai SQ, Liu JC. Notch signaling and non-small cell lung cancer. *ONCOL LETT.* 2018;15:3415–21.
50. Henneke I, Greschus S, Savai R, Korfei M, Markart P, Mahavadi P, et al. Inhibition of urokinase activity reduces primary tumor growth and metastasis formation in a murine lung carcinoma model. *Am J Respir Crit Care Med.* 2010;181:611–9.
51. Mauro CD, Pesapane A, Formisano L, Rosa R, D'Amato V, Ciciola P, et al. Urokinase-type plasminogen activator receptor (uPAR) expression enhances invasion and metastasis in RAS mutated tumors. *Sci*

52. Kitazawa S, Nishizawa S, Nakagawa H, Funata M, Nishimura K, Soga T, et al. Cancer with low cathepsin D levels is susceptible to vacuolar (H(+))-ATPase inhibition. *CANCER SCI.* 2017;108:1185–93.
53. Coelho M, Imperatori A, Chiaravalli AM, Franzi F, Castiglioni M, Rasini E, et al. Beta1- and Beta2-Adrenoceptors Expression Patterns in Human Non-small Cell Lung Cancer: Relationship with Cancer Histology. *J Neuroimmune Pharmacol.* 2019;14:697–708.
54. Huang Q, Tan Q, Mao K, Yang G, Ma G, Luo P, et al. The role of adrenergic receptors in lung cancer. *AM J CANCER RES.* 2018;8:2227–37.
55. Knizhnik AV, Kovaleva OB, Laktionov KK, Mochal'Nikova VV, Komel'Kov AV, Chevkina EM, et al. Arf6, RalA and BIRC5 protein expression in non small cell lung cancer. 2011.
56. Li F, Aljahdali I, Ling X. Cancer therapeutics using survivin BIRC5 as a target: what can we do after over two decades of study? *J Exp Clin Cancer Res.* 2019;38:368.
57. Kuang Y, Guo W, Ling J, Xu D, Liao Y, Zhao H, et al. Iron-dependent CDK1 activity promotes lung carcinogenesis via activation of the GP130/STAT3 signaling pathway. *CELL DEATH DIS.* 2019;10:297.
58. Wang S, Sun H, Zhan X, Wang Q. MicroRNA718 serves a tumorsuppressive role in nonsmall cell lung cancer by directly targeting CCNB1. *INT J MOL MED.* 2020;45:33–44.
59. Ruan JS, Zhou H, Yang L, Wang L, Zong SJ, Wang SM. CCNA2 facilitates epithelial-to-mesenchymal transition via the integrin $\alpha\beta3$ signaling in NSCLC. *Int J Clin Exp Pathol.* 2017;8:8324–33.
60. Wang M, Xie S, Yuan W, Xie T, Jamal M, Huang J, et al. Minichromosome maintenance protein 10 as a marker for proliferation and prognosis in lung cancer. *INT J ONCOL.* 2019;55:1349–60.
61. Rundle S, Bradbury A, Drew Y, Curtin NJ. Targeting the ATR-CHK1 Axis in Cancer Therapy. *Cancers (Basel).* 2017;9: .
62. Oo ZY, Proctor M, Stevenson AJ, Nazareth D, Fernando M, Daignault SM, et al. Combined use of subclinical hydroxyurea and CHK1 inhibitor effectively controls melanoma and lung cancer progression, with reduced normal tissue toxicity compared to gemcitabine. *MOL ONCOL.* 2019;13:1503–18.
63. Wojtukiewicz MZ, Hempel D, Sierko E, Tucker SC, Honn KV. Thrombin-unique coagulation system protein with multifaceted impacts on cancer and metastasis. *Cancer Metastasis Rev.* 2016;35:213–33.
64. Nadir Y. Decreasing Tumor Growth and Angiogenesis by Inhibition of Coagulation. *SEMIN THROMB HEMOST.* 2019;45:622–8.
65. Wang J, Wu Y, Guo J, Fei X, Yu L, Ma S. Adipocyte-derived exosomes promote lung cancer metastasis by increasing MMP9 activity via transferring MMP3 to lung cancer cells. *Oncotarget.* 2017;8:81880–91.

Figures

Figure 1

C-T network construction of Shenlian Capsule. A: 11 Herbs Included in Shenlian Capsules. B: Number of compounds in each herbal medicine in Shenlian capsule. C: The number of compounds remaining in the each index of ADME. D: Venn diagram of the

components of Shenlian capsule filtered by each index of ADME. E: Potential Compounds-Target (C-T) Networks of Shenlian Capsules.

Figure 2

C-D-T network construction of Shenlian Capsule, enrichment analysis of network targets and analysis of KEGG pathway. A: Four databases for non-small cell lung cancer targets. B: Non-small cell lung cancer target(D-T) network. D-T network consists of 226 NSCLC-related targets. C: C-T network and D-T network intersect to get the Venn diagram of 47 drug-acting NSCLC targets. D: Drug Compound Target Disease-C-D-T Network. The C-D-T network consists of 141 nodes and 541 edges. The nodes are composed of 47 NSCLC related targets and 94 potential compound nodes. E: Analysis of target protein GO enrichment based on C-D-T network. F: Analysis of target protein KEGG enrichment based on C-D-T network.

Figure 3

PPI Network core targets evaluation and Cluster analysis. A: PPI network construction results. (a) PPI network construction result representing potential compounds acting as targets for Shenlian capsules. The network includes 8422 targets and 186422 edges. (b) PPI network construction result representing NSCLC disease-related targets. The network consists of 6554 targets and 160234 edges. B: The core targets were identified by intersecting the C-T PPI network and the D-

T PPI network.(a)Represents a schematic diagram of a C-D-T PPI network. (b) Primary screening of core targets using Degree> 62 for C-D-T PPI network. (c) Topological indicators were used to screen the core targets of the C-D-T PPI network. C:Cluster analysis of core target PPI network to obtain 4 clusters. cluster1 consists of 22 nodes and 229 edges, cluster2 consists of 36 nodes and 296 edges, cluster3 consists of 30 nodes and 183 edges, cluster4 consists of 10 nodes and 16 edges.

Figure 4

Enrichment analysis of GO and KEGG pathway for each cluster. A: Cluster analysis of core target PPI network. B: Go enrichment analysis results for each cluster. C: KEGG pathway analysis results for each cluster.

Figure 5

Effect of Shenlian capsule on the role of LUSC survival-related gene. A: (a, b) Heat map and volcano map of differential gene expression in LUSC disease. (c) The Venn diagram of the intersection of LUSC disease survival-related genes with the target of Shenlian Capsule, obtained the active compounds of Shenlian Capsule that affected two survival-related genes about LUSC. B: (a) It representing the

difference in expression of the CTSD genes in normal tissues and cancer tissues. (b) It represents the survival curve of high and low expression of CTSD gene. (c) It represents the difference in expression of PLAU gene between normal tissues and cancer tissues. (d) It represents the survival curve of high and low expression of PLAU gene. C:(a b) The two-dimensional structure and three-dimensional structure diagrams show the lowest energy level conformation of the active compound of MOL000098 docked with the CTSD protein target. (c d) The two-dimensional structure and three-dimensional structure diagrams show the lowest energy level conformation of the active compound of MOL000449 docked with the PLUA protein target.

Figure 6

Effect of Shenlian capsule on the role of LUAD survival-related gene A:(a, b) Heat map and volcano map of differential gene expression in LUAD disease. (c) The Venn diagram of the intersection of LUAD disease survival-related genes with the target of Shenlian Capsule obtained the active compounds of Shenlian Capsule that affected eight survival-related genes about LUAD. B: Represents the difference in Loading [MathJax]/jax/output/CommonHTML/fonts/TeX/fontdata.js eight LUAD survival-related genes. C: Survival

curves representing the high expression group and low expression group of eight survival-related genes about LUAD. D: Two-dimensional structure diagram of active compounds after docking with survival-related genes. E: Three-dimensional structure diagram of active compounds after docking with survival-related genes.

Figure 7

Active compounds of Shenlian Capsules act on NSCLC survival-related protein targets.

Supplementary Files

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

- [Additionalfile5.xlsx](#)
- [Additionalfile4.xlsx](#)
- [Additionalfile3.xlsx](#)
- [Additionalfile2.xlsx](#)
- [Additionalfile1.xlsx](#)