The molecular mechanism of TiaoGanYiPi formula for treatment of chronic hepatitis B by network pharmacology and molecular docking verification

Xu Cao

Dongzhimen Hospital, Beijing University of Chinese Medicine

Xiaobin Zao

Dongzhimen Hospital, Beijing University of Chinese Medicine

Baiquan Xue

The First People's Hospital of Jinzhou District

Hening Chen

Dongzhimen Hospital, Beijing University of Chinese Medicine

Jiaxin Zhang

Dongzhimen Hospital, Beijing University of Chinese Medicine

Shuo Li

Dongzhimen Hospital, Beijing University of Chinese Medicine

Xiaobin Li

Dongzhimen Hospital, Beijing University of Chinese Medicine

Shun Zhu

Dongzhimen Hospital, Beijing University of Chinese Medicine

Rui Guo

Beijing University of Chinese Medicine

Xiaoke Li

Dongzhimen Hospital, Beijing University of Chinese Medicine

Yongan Ye (✉ yeyongan@vip.163.com)

Dongzhimen Hospital, Beijing University of Chinese Medicine

Research Article

Keywords: TiaoGanYiPi formula, network pharmacology, molecular docking, chronic hepatitis B, bioinformatics

Posted Date: December 15th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-120141/v1>

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

[Read Full License](#)

Version of Record: A version of this preprint was published at Scientific Reports on April 16th, 2021. See the published version at <https://doi.org/10.1038/s41598-021-87812-9>.

Abstract

The Chinese herbal formula Tiao-Gan-Yi-Pi (TGYP) showed effective against Chronic Hepatitis B (CHB). In this study, we aimed to clarify the mechanisms and potential targets between TGYP and CHB through network pharmacology and molecular docking verification. The compounds of TGYP were identified in the TCMSP and CNKI databases, and their putative targets were predicted through SwissTargetPrediction and STITCH databases. The targets of CHB were obtained from the GeneCards, NCBI Gene, and DisGeNET databases. The above mentioned data were visualized using Cytoscape, and molecular docking showed the relationship between them. The expression of key targets was verified in GEO databases. Hence, we screened out 11 TGYP-related key targets for CHB included ABL1, CASP8, CCNA2, CCNB1, CDK4, CDKN1A, EP300, HIF1A, IGF1R, MAP2K1 and PGR. The key targets were predominantly enriched in the cancer, cell cycle and hepatitis B pathways and involved in the positive regulation of fibroblast proliferation, signal transduction, and negative regulation of gene expression biological processes, and expression of key target genes was related to HBV infection and liver inflammation. Through this newly constructed interaction network between TGYP and CHB, we identified active compounds and targets which could be further used for providing clinical guidance.

1. Introduction

It has been a global public health problem for the Hepatitis B virus (HBV) infection for a long time¹. Worldwide estimates suggested that more than 2 billion people have been infected with HBV, with 248 million of these people are chronic hepatitis B (CHB) patients, and about 15% ~ 25% of CHB patients died from cirrhosis or liver cancer², which caused a substantial global burden³. In China, there are 20 ~ 30 million CHB patients, and the mortality rate of HBV-related complications is higher than that in the world⁴. The disease process of CHB is affected by both HBV and the host immune system, while the pathogenesis of CHB is complex and still unclear. Meanwhile, Existing treatment options [nucleos(t)ide analogs (NAs) and interferons (IFN)] are limited in their efficacy and scope of application. Studies have suggested that discontinuation of NAs lacked operability⁵, whilst oral antiviral drugs had limitations in improving complications⁶. Meanwhile, the contraindications and side effects hindered the clinical application of IFN⁷.

For a long time, Chinese herbal medicines (CHMs) play a crucial role in CHB treatment and show superiority^{8,9}. Among CHMs, the Tiao-gan-yi-pi formula (TGYP) was reported with efficacy for CHB treatment, as reported by our previous multi-center, double-blind, randomized placebo-controlled trial¹⁰. In Chinese Medicine theory, stagnation of Gan and deficiency of Pi qi, and damp-heat in the Gan and Dan, is a compound core syndrome of chronic hepatitis B¹¹. The TGYP is targeting at the core syndrome of CHB, which reinforces and motivate qi, and dispel damp-heat. Prescription components may play a role in the various multi-targeting and synergistic effects of the CHM formula. However, the pharmacological underlying curative mechanisms of TGYP regarding the therapeutic effect of CHB have not been fully

explained. Here, we aimed to explore the relationship between CHB and TGYP, identify the targets and molecular mechanisms of TGYP on CHB.

In this study, we performed network pharmacology based on system biology and molecular docking, screened out 11 key target genes that exert essential functions during the treatment of TGYP on CHB. These results revealed the relationship between TGYP and CHB, indicating directions for further research.

2. Results

2.1 Workflow of TGYP

Based on the previous work, which showed the anti-HBV effect of TGYP, we designed the flowchart of this study (Fig. 1). We aimed to determine the key target genes and regulatory mechanism of TGYP treatment against CHB through network pharmacology and database mining methods.

2.2 Compound-Putative Targets Network

We first analyzed the four main herbs in TGYP by CNKI and TCMSP databases and found 26 active components in Chaihu, 25 active components in Huangqi, 25 active components in Danggui, and 17 active components in Kudiding, respectively (Figure S1A). Among them, Chaihu and Huangqi have three common components, Chaihu and Danggui have two common components which were showed in Table 1, the detailed information of all components was shown in Table S1. Next, according to the active components, we identified 775 candidate target genes by SwissTargetPrediction and 1,198 candidate target genes by STITCH analysis for TGYP (Table S2). After eliminating the redundancy, a total of 1,764 putative targets were collected (Figure S1B) and visualized as the compound-putative target network by Cytoscape (Figure S1C). The compound-putative target network showed 1,856 nodes and 4,941 edges. According to combine degree, the top three ranked compounds were Isorhamnetin (Degree = 147), Quercetin (Degree = 145), and Kaempferol (Degree = 142), indicated their vital roles in TGYP and defined as key compounds.

Table 1
Common components information of TGYF

Molecule name	Abbreviation	Canonical SMILES	Stem from herb
quercetin	A1	<chem>C1=CC(=C(C=C1C2=C(C(=O)C3=C(C=C(C(=C3O2)O)O)O)O)O</chem>	Chaihu and Huangqi
isorhamnetin	A2	<chem>COC1=C(C=CC(=C1)C2=C(C(=O)C3=C(C=C(C(=C3O2)O)O)O)O</chem>	Chaihu and Huangqi
kaempferol	A3	<chem>C1=CC(=CC=C1C2=C(C(=O)C3=C(C=C(C(=C3O2)O)O)O)O</chem>	Chaihu and Huangqi
Stigmasterol	B1	<chem>CCC(C=CC(C)C1CCC2C1(CCC3C2CC=C4C3(CCC(C4)O)C)C)C(C)C</chem>	Chaihu and Danggui
Psoralen	B2	<chem>C1=CC(=O)OC2=CC3=C(C=CO3)C=C21</chem>	Chaihu and Danggui

2.3 Functional enrichment of TGYF target genes

To explore the biological function of TGYF target genes, the GO and KEGG pathway enrichment analysis were performed by DAVID. The GO enrichment analysis revealed that the GO-CC of TGYF target genes were predominantly involved in the cytoplasm, plasma membrane and nucleus (Fig. 2A); the GO-MF of TGYF target genes mainly included protein binding, ATP binding, and protein homo-dimerization activity (Fig. 2B); and the GO-BP of TGYF target genes concentrated primarily on signal transduction, oxidation-reduction process, and positive regulation of transcription from RNA polymerase II promoter (Fig. 2C). Meanwhile, the KEGG enrichment analysis suggested that TGYF target genes were primarily associated with metabolic, neuroactive ligand-receptor interactions and cancer pathways (Fig. 2D). The results above suggest that TGYF target genes are enriched in viral carcinogenesis and hepatitis B pathways closely related to CHB, and the GO results broadly correspond to this. However, the pathway analysis also enriched in metabolic, neuroactive ligand-receptor interaction, calcium signaling pathway, and more, suggesting that the TGYF formula has a wide range of action pathways and complex mechanisms that are consistent with the properties of traditional Chinese medicine compound in treating diseases. Of course, partial pathways that do not appear to be closely related to CHB still need to be further explored.

2.4 Identification and PPI Network Construction of OGEs

We analyzed and screened CHB disease targets, where we identified 8,037 targets in the GeneCards database, 415 targets in the DisGeNET database, and 236 targets in the NCBI Gene database, respectively (Table S3), and subsequently obtained 8,083 disease targets of CHB after eliminating the redundancy and taking the union of these three sets (Figure S2A). Next, we examined the intersection of TGYF and

CHB target genes and obtained 1,200 OGEs from the Venn diagrams tool (Figure S2B). After that, we constructed a PPI network of OGEs and analyzed their corresponding secondary proteins from the STRING tool (Figure S2C). The PPI network results showed the interactions between OGEs, and there were highly tight connections existed which indicated key target genes.

2.5 Identification of the key target genes through modules analysis

Since one gene lacked a corresponding protein structure, and while seven proteins did not interact with others based on PPI network, the remaining 1,192 proteins from STRING data were imported into the Cytoscape to further obtain interaction networks (Fig. 3A). Filtering with the network topological parameters, when the degree is greater than twice the median (degree>78), we get the preliminary hub network (Fig. 3B). Based on the fact that BC, CC, DC, LAC, NC, and SC are above the median (BC>0.001, CC>0.574, DC>61, LAC>38.452, NC>89.011, and SC>5.964E + 35), we identified 44 highly connected nodes as hub networks (Fig. 3C). The MCODE plugin in Cytoscape was used for a module analysis of the targets in the hub network. The module analysis identified four clustering modules, and we chose the key module, which had the highest network heterogeneity and network centralization at 0.242 and 0.333, respectively. The clustering key module we obtained represents a key target for treating CHB with TGYP. We acquired a dense region network of 11 key targets (Fig. 3D), including ABL1, CASP8, CCNA2, CCNB1, CDK4, CDKN1A, EP300, HIF1A, IGF1R, MAP2K1, and PGR.

2.6 Functional enrichment analysis of 11 key target genes

Functional enrichment was performed after data screening to further understand the biological behaviors of key targets (Table S4). The GO-CC analysis showed the main targets to be mainly enriched in the nucleoplasm, nucleus, and cytosol (Fig. 4A). GO-MF analysis suggested that the key targets were strongly associated with protein binding, protein kinase binding, and protein tyrosine kinase activity (Fig. 4B). And the key targets were significantly related to multiple GO-BP, including positive regulation of fibroblast proliferation, signal transduction, and negative regulation of gene expression (Fig. 4C). The pathway enrichment results suggested that the key targets were mostly involved in pathways of cancer, cell cycle, hepatitis B and others (Fig. 4D). Based on the aforementioned results of GO-BP and KEGG, TGYP might have an effect on CHB related complications, such as liver fibrosis and liver cancer, which require further exploration. In general, these biological processes and signaling pathways may be linked to beneficial effects of TGYP against CHB.

2.7 The molecular docking of key compounds and key target genes

To better understand the interactions between key compounds and key target genes predicted by SwissTargetPrediction and STITCH, we constructed an interaction network between key target genes and corresponding active compounds in TGYP (Fig. 5A). According to the compound-putative target network analysis results, the key compounds were Isorhamnetin, Quercetin, and Kaempferol. Their 2D structures were shown in Fig. 5B. To further explore the interactions between key compounds and key target genes, we performed molecular docking by AutoDock Vina software (Fig. 5C). The docking capabilities of the protein-compounds are shown in Table 2. The docking simulations results included key target genes

docked with key compounds and the most probable active compounds, respectively. The results further clarified the regulation mechanism of TGYP on the key targets on the protein level.

Table 2
Functional enrichment of key targets.

Targets	PBD ID	Compounds	Herb	Affinity (kcal/mol)
ABL1	6BL8	(+)-Anomalin	Radix Bupleuri	-8.4
CASP8	6PX9	Baicalin	Radix Bupleuri	-8.1
CASP8	6PX9	Cubebin	Radix Bupleuri	-7.1
CCNA2	1E9H	corycavine	Corydalis Bungeanae Herba	-10.1
CCNB1	6GU2	Calycosin	Hedysarum Multijugum Maxim	-7.9
CCNB1	6GU2	quercetin	Radix Bupleuri and Hedysarum Multijugum Maxim	-7.5
CCNB1	6GU2	Isorhamnetin	Radix Bupleuri and Hedysarum Multijugum Maxim	-7.6
CCNB1	6GU2	Kaempferol	Radix Bupleuri and Hedysarum Multijugum Maxim	-8.0
CDK4	5FWP	(+)-Anomalin	Radix Bupleuri	-8.3
CDKN1A	6P8H	Calycosin	Hedysarum Multijugum Maxim	-6.6
EP300	6PF1	(3R)-3-(2-hydroxy-3,4-dimethoxyphenyl)chroman-7-ol	Hedysarum Multijugum Maxim	-8.2
HIF1A	3HQU	quercetin	Radix Bupleuri and Hedysarum Multijugum Maxim	-7.7
HIF1A	3HQU	Saikogenin F	Radix Bupleuri	-6.8
IGF1R	3D94	quercetin	Radix Bupleuri and Hedysarum Multijugum Maxim	-8.3
IGF1R	3D94	kaempferol	Radix Bupleuri and Hedysarum Multijugum Maxim	-8.2

Targets	PBD ID	Compounds	Herb	Affinity (kcal/mol)
IGF1R	3D94	Isorhamnetin	Radix Bupleuri and Hedysarum Multijugum Maxim	-8.1
MAP2K1	3EQF	8-oxocorynoline	Corydalis Bungeanae Herba	-12.7
PGR	1A28	Secoisolariciresinol diglucoside	Radix Bupleuri	-8.1
PGR	1A28	hederagenin	Hedysarum Multijugum Maxim	-7.4

2.8 Expression of key genes in CHB and correlation with liver inflammation

We analyzed the key target genes' expression and their correlations with liver inflammation in the GEO dataset GSE83148. Compared to healthy controls, some key target genes were significantly higher expressed in CHB patients ($P < 0.05$), including ABL1, CASP8, CCNA2, CCNB1, CDK4, CDKN1A, HIF1A, and IGF1R (Fig. 6A). Results suggested the expression of these key genes were possibly related to the disease status of CHB. Furthermore, CHB patients with significant liver inflammation could benefit from improved antiviral effect¹². To further explore the relationship between key target genes and liver inflammation in CHB patients, we compared the expression of 71 patients with abnormal serum ALT or AST levels ($ALT \geq 40$ or $AST \geq 35$) and 34 patients with normal serum ALT and AST levels ($ALT < 40$ and $AST < 35$). Results showed that expression of ABL1, CCNA2, CCNB1, CDKN1A, and HIF1A was significantly higher in the abnormal group than in the normal group, whereas MAP2K1 and PGR were reversed (Fig. 6B). These results indicate that expression of these key genes might be related to inflammatory CHB. TGYP might alter the disease status and the inflammatory state of CHB by regulating the expression of these genes.

3. Discussion

CHB is a refractory disease which poses great threats to public health, especially in the Asia-Pacific region. It is closely related to the occurrence and development of liver fibrosis and liver cancer¹³. At present, NAs and IFNs are the first options for treatment, whereas both of them have shortcomings. NAs cannot eliminate HBV cccDNA in hepatocytes to achieve complete cure, while IFNs has many contraindications and side effects¹⁴. In contrast, traditional Chinese medicines have unique advantages against CHB, such as better safety profile and lower price⁸. Our preliminary study confirmed the capability and safety of TGYP for improving HBeAg loss after combination with entecavir (ETV), thereby reducing disease-related risks in long terms¹⁰. For TGYP formula, the main components are Chaihu, Diding, Huangqi and Danggui. It is noteworthy that other studies have reported that Huangqi could relieve HBV induced hepatitis by inhibiting NF- κ B signaling¹⁵, and Chaihu could promote HBeAg clearance by

inhibiting HBV replication^{16,17}. Combined these studies, it suggested that multiple active compounds of TGYP could have inhibitory effect on HBV in different ways. It is necessary to explore internal mechanisms to find more treatment targets.

To further determine the mechanism of TGYP on CHB, we performed network pharmacology and database mining in this study: we firstly screened out TGYP target genes based on drug active compounds of the four main herbs. Among the active molecules, the top three active compounds with the highest network connectivity were Isorhamnetin, Quercetin, and Kaempferol. These compounds were observed inhibiting liver inflammation or anti-HBV effects, including that Isorhamnetin could significantly reduce serum levels of liver enzymes and pro-inflammatory cytokines in rats¹⁸; Quercetin could strongly inhibit HBeAg synthesis and slightly anti-HBV surface antigen activity¹⁹, and Kaempferol could significantly reduce TNF- α expression and inflammatory responses²⁰. These results suggested that TGYP may exert anti-HBV function via these active compounds and their target genes and pathways. We then combined CHB related genes to construct the interaction network of TGYP and CHB, where 11 key target genes were identified. Molecular docking analyses predicted direct binding between core active compounds and key target genes. We also found that in the GSE83148 database, compared with healthy controls, the expression of key target genes including ABL1, CASP8, CCNA2, CCNB1, CDK4, CDKN1A, HIF1A, and IGF1R were all up-regulated in liver tissues of CHB patients. Thus, the core active compounds of TGYP might exert functions by regulating the expression and protein activities of the key target genes.

Additionally, we found that the key target genes were mainly enriched in cancer, cell cycle, and hepatitis B signaling pathways using the KEGG pathway enrichment analysis. Specifically, among the 11 key target genes, CCNB1, CDKN1A, EP300, ABL1, CDK4, and CCNA2 are enriched in the cell cycle pathway, and previous studies have shown that the expression of these factors could regulate the cell cycle to varying degrees, promote or inhibit cell proliferation^{21–26}. Hepatocyte proliferation has been reported to accelerate during HBV infection²⁷, leading to the downregulation of NTCP in hepatocytes, which contributes to HBV cccDNA loss and further accelerates HBV clearance^{12,28}. Our study suggested that TGYP might exert anti-HBV effects by regulating the cell-cycle-related genes which manipulate hepatocyte proliferation. However, the specific regulatory mechanism needs to be further investigated in HBV replication models.

Key target genes enriched in the hepatitis B pathway include MAP2K1 and CASP8. For MAP2K1, studies have shown that activation of the Ras-Raf-MAP2K1-ERK pathway could inhibit HBV replication^{29,30}, while HBV infection could induce IL-23 and IL-6 expression in hepatocytes through MAP2K1, which could inhibit anti-HBV immune response^{31,32}. In addition, HBx could induce hepatocyte apoptosis by activating the MAPK pathway, which was also an important cause of CHB³³. For CASP8, its expression on peripheral blood mononuclear cells was positively correlated with HBV DNA level in CHB patients³⁴. The combination of Caspase-8 and TNFAIP8L2 could inhibit the activities of AP-1 and NF- κ B, which in turn promoted Fas-induced apoptosis of immune cells³⁵. Moreover, HBx could increase the expression level of caspase-8 and promote apoptosis induced by exogenous pathways mediated by death receptors³⁶.

Meanwhile, studies have shown that HBV infection could inhibit the innate immune pathways in hepatocytes and the liver's adaptive immune system, which was critically important for the development of CHB³⁷. In the current study, we found the expression of multiple target genes in liver tissue was correlated with the severity of hepatocyte inflammation. It is hypothesized that the TGYP might slow down the progression of CHB by regulating the immune response in the liver, which needs to be further investigated in animal models.

Finally, IGF1R and HIF1A genes enriched in cancer pathways were closely related to the occurrence and development of HBV-related HCC. Studies have shown that HBx could trans-activate endogenous IGF1R expression and promote cell mitosis, thereby promoting the proliferation of liver cancer cells³⁸. Meanwhile, HBx could enhance the transcriptional activity of HIF1A, thus promoting para-tumor angiogenesis^{39,40}. IGF1R and HIF1A promote the occurrence and metastasis of HBV-associated liver cancer. Our results suggest that TGYP might affect the expression and activation of these genes, thereby playing a role in the inhibition of HBV-related HCC. Recent studies have shown that although NAs may inhibit HBV replication in hepatocytes, they cannot completely reverse HBV-induced signal transduction pathway activation or cell cycle regulatory protein change⁴¹. Those infected hepatocytes were still in transition to cancer cells. In contrast, our preliminary study demonstrated the superiority of integrative therapy regarding HBeAg clearance, the above results suggested TGYP may exert beneficial to prevent the risk of HCC in CHB patients. These results need to be validated in a larger sample, whilst the mechanisms are to be explored with *in vitro* and *in vivo* models.

In this study, a network was constructed between active compounds of TGYP and CHB related genes, demonstrating that this CHM formular might exert an anti-HBV effect through regulating hepatocyte proliferation, liver inflammation, and immune response. These findings could provide further insight into the mechanisms of CHM-based CHB therapy and assist in screening potential therapeutic targets in the future.

4. Methods

4.1 Collection of chemical compounds in TGYP.

Compounds of the main herbs in TGYP, including Chaihu, Huangqi, Danggui, and Kudiding were searched in CNKI and TCMSP database (<http://tcmssp.com/tcmssp.php>)⁴². The active compounds in TCMSP were mainly filtered by integrating the pharmacokinetic properties comprising oral bioavailability (OB) $\geq 30\%$ ⁴³, drug-likeness (DL) ≥ 0.18 ⁴⁴, as well as half-life (HL) ≥ 4 h. The corresponding compounds chemical 2D structure SDF and canonical SMILES acquired from PubChem database (<https://pubchem.ncbi.nlm.nih.gov>)⁴⁵. Data formats were transformed with Open Babel GUI software⁴⁶. The Venn diagram was charted by the Venn tool (<http://bioinformatics.psb.ugent.be/webtools/Venn/>).

4.2 Prediction of putative targets in TGYP.

The canonical SMILES or SDF of the active compounds were imported into the SwissTargetPrediction (<http://www.swisstargetprediction.ch/>)⁴⁷ and the STITCH databases (<http://stitch.embl.de/>)⁴⁸, respectively. Putative targets of TGYF were collected under the following conditions: selecting the species is homo sapiens. The probability value of each potential target listed in the SwissTargetPrediction database was used to investigate the accuracy of the current predictions, whose probability value ≥ 0.1 was identified. In addition, the potential target proteins with confidence score ≥ 1.5 in the STITCH database.

4.3 Identification of CHB Targets

The target genes related to CHB were gathered from followed three databases. The GeneCards database (<https://www.genecards.org/>)⁴⁹, the DisGeNET database (<https://www.disgenet.org/>)⁵⁰ and the NCBI gene database (<https://www.ncbi.nlm.nih.gov/gene/>)⁵¹. The search condition was using the keyword "Chronic Hepatitis B" and selecting the organisms "Homo sapiens" from the three platforms. The above gene ID was identified and standardized in the UniProt database (<https://www.uniprot.org/>)⁵². We used Retrieve/ID mapping tool of UniProt to convert identifiers, which are of a different type to UniProt identifiers.

4.4 Network construction

The overlapped genes (OGEs) of the compound-putative targets and disease targets were charted by the Venn tool. We inputted the OGEs into the STRING database (<https://string-db.org/>)⁵³ with confidence scores ≥ 0.4 and the species limited to "Homo sapiens", and exported protein-protein interaction (PPI) data. Analyzed the PPI data and constructed network by the Cytoscape 3.8.0 (<https://cytoscape.org/>)⁵⁴, and the non-connection genes were removed. We use the Analyzer plugin to analyze the PPI network and get the degree, through taking the over a double median of degree to get the preliminary hub network⁵⁵. Then, using the CytoNCA plugin to analyze the preliminary hub network to get network topological parameters: Betweenness Connectivity (BC), Closeness Connectivity (CC), Degree Connectivity (DC), Local Average Connectivity (LAC), Neighbor Connectivity (NC) and Subgraph Centrality (SC). We took the excess median of BC, CC, DC, LAC, NC, and SC to obtain hub network⁵⁶. The MCODE application calculated the hub network to get the dense regions⁵⁷. The genes in dense regions from MCODE is called key targets.

4.5 Analysis of GO and KEGG Pathway

The screened out genes were analyzed by the Gene Ontology (GO) function analysis included Biological Process (BP), Molecular Function (MF) and Cellular Component (CC)⁵⁸, and the Kyoto Encyclopedia of Gene and Genome (KEGG)

analysis (<https://www.genome.jp/kegg/>)⁵⁹. The GO function enrichment and KEGG pathway enrichment were analyzed by DAVID database (<https://david.ncifcrf.gov/>)⁶⁰. The filtering of retrieval results is with a threshold value of $p < 0.05$ and count in descending order. The value of P reflected the significance of protein biological function, and the purpose of filtering is to select the most enriched biological

annotation from a given list of genes. Bubble Diagrams were drawn by R software with package ggplot2 (<https://cran.r-project.org/web/packages/ggplot2/>).

4.6 Molecular Docking Simulation

Molecular docking simulation was performed to verify the critical components' binding ability to the key targets and explore their accurate binding modes. The macromolecular protein target receptors were acquired from the RCSB PDB database (<https://www.rcsb.org>)^{61 62}. Moreover, the 2D structures of the compounds were obtained from the PubChem Database. AutoDock Vina could markedly improve the average accuracy of the combined mode prediction⁶³. Thereby the Molecular docking simulation was performed by AutoDock Vina software, and the result was visualized by PyMOL software⁶⁴. In this analysis, the value of the vina score indicates the binding activity between a compound and a protein, and the lower the vina score, the more stably the compound binds to the protein.

4.7 Verification the expression of key genes

The expression of key genes was verified in the GEO database (<http://www.ncbi.nlm.gov/geo/>)⁶⁵. The keywords were set as "hepatitis B" and "homo sapiens", the tissue type was limited to liver tissue, and the gene expression data of GSE83148⁶⁶ was downloaded from the GEO database. The platform of dataset GSE83148 is the GPL570 (HG-U133_Plus_2) Affymetrix Human Genome U133 Plus 2.0 Array, which includes 122 patients with CHB and 6 healthy liver samples, and includes 71 patients with abnormal transaminase and 34 patients with normal transaminase, which threshold of alanine aminotransferase (ALT) was 40 and the threshold of aspartate aminotransferase (AST) was 35.

Abbreviations

Tiao-Gan-Yi-Pi (TGYP); Chronic Hepatitis B (CHB); overlapped genes (OGEs); Hepatitis B virus (HBV); nucleos(t)ide analogs (NAs); interferons (IFN); Chinese herbal medicines (CHMs); oral bioavailability (OB); drug-likeness (DL); half-life (HL); protein-protein interaction (PPI); Betweenness Connectivity (BC); Closeness Connectivity (CC); Degree Connectivity (DC); Local Average Connectivity (LAC); Neighbor Connectivity (NC); Subgraph Centrality (SC); Gene Ontology (GO); Biological Process (BP); Molecular Function (MF); Cellular Component (CC); Kyoto Encyclopedia of Gene and Genome (KEGG); alanine aminotransferase (ALT); aspartate aminotransferase (AST); entecavir (ETV); ABL Proto-Oncogene 1, Non-Receptor Tyrosine Kinase (ABL1); Caspase 8 (CASP8); Cyclin A2 (CCNA2); Cyclin B1 (CCNB1); Cyclin Dependent Kinase 4 (CDK4); Cyclin Dependent Kinase Inhibitor 1A (CDKN1A); E1A Binding Protein P300 (EP300); Hypoxia Inducible Factor 1 Subunit Alpha (HIF1A); Insulin Like Growth Factor 1 Receptor (IGF1R); Mitogen-Activated Protein Kinase Kinase 1 (MAP2K1); Progesterone Receptor (PGR)

Declarations

Data availability

All the relevant data is provided within the paper and its supporting information files. The datasets analyzed during the current study available from the corresponding author on reasonable request.

Data statement

This work described was original research that has not been published previously, and not under consideration for publication elsewhere, in whole or in part.

Acknowledgements

This research was supported by the China National Science and Technology major projects, 13th 5-year plan, [grant number 2018ZX10725505] and National Nature Science Foundation of China [grant number 81804033].

Authors Contributions

C.X. and Z.X.B. wrote the manuscript, contributed equally to this work. Y.Y.A., L.X.K. and C.X. initiated and designed the project. X.B.Q., C.H.N., L.S. and G.R. collected the data. Z.X.B., C.X. and Z.J.X. performed the data analyses. L.X.B. and Z.S. prepared Figs. 1–8. All the authors were responsible for reviewing the data. In addition, all the authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

References

- 1 Nelson, N. P., Easterbrook, P. J. & McMahon, B. J. Epidemiology of Hepatitis B Virus Infection and Impact of Vaccination on Disease. *Clin Liver Dis* **20**, 607-628, doi:10.1016/j.cld.2016.06.006 (2016).
- 2 Schweitzer, A., Horn, J., Mikolajczyk, R., Krause, G. & Ott, J. Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013. *Lancet (London, England)* **386**, 1546-1555, doi:10.1016/s0140-6736(15)61412-x (2015).
- 3 The global, regional, and national burden of cirrhosis by cause in 195 countries and territories, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *The lancet. Gastroenterology & hepatology* **5**, 245-266, doi:10.1016/s2468-1253(19)30349-8 (2020).
- 4 Liu, J., Liang, W., Jing, W. & Liu, M. Countdown to 2030: eliminating hepatitis B disease, China. *Bulletin of the World Health Organization* **97**, 230-238, doi:10.2471/blt.18.219469 (2019).
- 5 Papatheodoridis, G. *et al.* Discontinuation of oral antivirals in chronic hepatitis B: A systematic review. *Hepatology (Baltimore, Md.)* **63**, 1481-1492, doi:10.1002/hep.28438 (2016).

- 6 Kim, T. *et al.* Hepatitis B virus DNA levels and overall survival in hepatitis B-related hepatocellular carcinoma patients with low-level viremia. *Journal of gastroenterology and hepatology* **34**, 2028-2035, doi:10.1111/jgh.14750 (2019).
- 7 [The guidelines of prevention and treatment for chronic hepatitis B (2019 version)]. *Zhonghua gan zang bing za zhi = Zhonghua ganzangbing zazhi = Chinese journal of hepatology* **27**, 938-961, doi:10.3760/cma.j.issn.1007-3418.2019.12.007 (2019).
- 8 Zhang, L. *et al.* Contemporary clinical research of traditional Chinese medicines for chronic hepatitis B in China: an analytical review. *Hepatology (Baltimore, Md.)* **51**, 690-698, doi:10.1002/hep.23384 (2010).
- 9 Chen, Y. & Zhu, J. Anti-HBV effect of individual traditional Chinese herbal medicine in vitro and in vivo: an analytic review. *J. Viral Hepat.* **20**, 445-452, doi:10.1111/jvh.12112 (2013).
- 10 Li, X. *et al.* Entecavir combining Chinese herbal medicine for HBeAg-positive chronic hepatitis B patients: a randomized, controlled trial. *Hepatology international*, doi:10.1007/s12072-020-10097-z (2020).
- 11 Chang, T. *et al.* Entecavir treatment for up to 5 years in patients with hepatitis B e antigen-positive chronic hepatitis B. *Hepatology (Baltimore, Md.)* **51**, 422-430, doi:10.1002/hep.23327 (2010).
- 12 Yan, Y. *et al.* Down-regulation of cell membrane localized NTCP expression in proliferating hepatocytes prevents hepatitis B virus infection. *Emerging microbes & infections* **8**, 879-894, doi:10.1080/22221751.2019.1625728 (2019).
- 13 Levrero, M. & Zucman-Rossi, J. Mechanisms of HBV-induced hepatocellular carcinoma. *Journal of hepatology* **64**, S84-s101, doi:10.1016/j.jhep.2016.02.021 (2016).
- 14 Nguyen, M. H., Wong, G., Gane, E., Kao, J. H. & Dusheiko, G. Hepatitis B Virus: Advances in Prevention, Diagnosis, and Therapy. *Clinical microbiology reviews* **33**, doi:10.1128/cmr.00046-19 (2020).
- 15 Wan, L. *et al.* Extracts of Qizhu decoction inhibit hepatitis and hepatocellular carcinoma in vitro and in C57BL/6 mice by suppressing NF- κ B signaling. *Scientific reports* **9**, 1415, doi:10.1038/s41598-018-38391-9 (2019).
- 16 Tajiri, H. *et al.* Effect of sho-saiko-to(xiao-chai-hu-tang) on HBeAg clearance in children with chronic hepatitis B virus infection and with sustained liver disease. *The American journal of Chinese medicine* **19**, 121-129, doi:10.1142/s0192415x91000193 (1991).
- 17 Chang, J. S. *et al.* Sho-saiko-to (Xiao-Chai-Hu-Tang) and crude saikosaponins inhibit hepatitis B virus in a stable HBV-producing cell line. *The American journal of Chinese medicine* **35**, 341-351, doi:10.1142/s0192415x07004862 (2007).

- 18 Lu, X. *et al.* Isorhamnetin: A hepatoprotective flavonoid inhibits apoptosis and autophagy via P38/PPAR- α pathway in mice. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie* **103**, 800-811, doi:10.1016/j.biopha.2018.04.016 (2018).
- 19 Lin, Y., Shen, C., Huang, Y. & Chang, Y. Homoflavonoids from *Ophioglossum petiolatum*. *Journal of natural products* **68**, 381-384, doi:10.1021/np0401819 (2005).
- 20 Cai, F. *et al.* Yinchenhao decoction suppresses rat liver fibrosis involved in an apoptosis regulation mechanism based on network pharmacology and transcriptomic analysis. *Biomedicine & pharmacotherapy = Biomedecine & pharmacotherapie* **114**, 108863, doi:10.1016/j.biopha.2019.108863 (2019).
- 21 Wang, Z. *et al.* Cyclin B1/Cdk1 coordinates mitochondrial respiration for cell-cycle G2/M progression. *Developmental cell* **29**, 217-232, doi:10.1016/j.devcel.2014.03.012 (2014).
- 22 Torres-Fernández, L. *et al.* The mRNA repressor TRIM71 cooperates with Nonsense-Mediated Decay factors to destabilize the mRNA of CDKN1A/p21. *Nucleic acids research* **47**, 11861-11879, doi:10.1093/nar/gkz1057 (2019).
- 23 Tsang, F. *et al.* Aberrant Super-Enhancer Landscape in Human Hepatocellular Carcinoma. *Hepatology (Baltimore, Md.)* **69**, 2502-2517, doi:10.1002/hep.30544 (2019).
- 24 Dasgupta, Y. *et al.* Normal ABL1 is a tumor suppressor and therapeutic target in human and mouse leukemias expressing oncogenic ABL1 kinases. *Blood* **127**, 2131-2143, doi:10.1182/blood-2015-11-681171 (2016).
- 25 Topacio, B. *et al.* Cyclin D-Cdk4,6 Drives Cell-Cycle Progression via the Retinoblastoma Protein's C-Terminal Helix. *Molecular cell* **74**, 758-770.e754, doi:10.1016/j.molcel.2019.03.020 (2019).
- 26 Hein, J. & Nilsson, J. Interphase APC/C-Cdc20 inhibition by cyclin A2-Cdk2 ensures efficient mitotic entry. *Nature communications* **7**, 10975, doi:10.1038/ncomms10975 (2016).
- 27 Guo, J. & Guo, H. Metabolism and function of hepatitis B virus cccDNA: Implications for the development of cccDNA-targeting antiviral therapeutics. *Antiviral research* **122**, 91-100, doi:10.1016/j.antiviral.2015.08.005 (2015).
- 28 Allweiss, L. *et al.* Proliferation of primary human hepatocytes and prevention of hepatitis B virus reinfection efficiently deplete nuclear cccDNA in vivo. *Gut* **67**, 542-552, doi:10.1136/gutjnl-2016-312162 (2018).
- 29 Zheng, Y., Li, J., Johnson, D. & Ou, J. Regulation of hepatitis B virus replication by the ras-mitogen-activated protein kinase signaling pathway. *Journal of virology* **77**, 7707-7712, doi:10.1128/jvi.77.14.7707-7712.2003 (2003).

- 30 Stöckl, L. *et al.* Integrity of c-Raf-1/MEK signal transduction cascade is essential for hepatitis B virus gene expression. *Oncogene* **22**, 2604-2610, doi:10.1038/sj.onc.1206320 (2003).
- 31 Xia, L. *et al.* Upregulation of IL-23 expression in patients with chronic hepatitis B is mediated by the HBx/ERK/NF- κ B pathway. *Journal of immunology (Baltimore, Md. : 1950)* **188**, 753-764, doi:10.4049/jimmunol.1101652 (2012).
- 32 Fang, Z. *et al.* Polarization of Monocytic Myeloid-Derived Suppressor Cells by Hepatitis B Surface Antigen Is Mediated via ERK/IL-6/STAT3 Signaling Feedback and Restrains the Activation of T Cells in Chronic Hepatitis B Virus Infection. *Journal of immunology (Baltimore, Md. : 1950)* **195**, 4873-4883, doi:10.4049/jimmunol.1501362 (2015).
- 33 Su, F. & Schneider, R. Hepatitis B virus HBx protein sensitizes cells to apoptotic killing by tumor necrosis factor alpha. *Proceedings of the National Academy of Sciences of the United States of America* **94**, 8744-8749, doi:10.1073/pnas.94.16.8744 (1997).
- 34 Guo, L. & Liang, Y. Expression of PBMC apoptosis-related factors in patients with chronic hepatitis B and their relationships with clinical prognosis. *Experimental and therapeutic medicine* **14**, 6007-6011, doi:10.3892/etm.2017.5299 (2017).
- 35 Ji, J., Zhang, Y. & Fan, Y. TIPE2 as a potential therapeutic target in chronic viral hepatitis. *Expert opinion on therapeutic targets* **23**, 485-493, doi:10.1080/14728222.2019.1608948 (2019).
- 36 Kuo, C. *et al.* Apoptosis induced by hepatitis B virus X protein in a CCL13-HBx stable cell line. *Oncology reports* **28**, 127-132, doi:10.3892/or.2012.1775 (2012).
- 37 Li, T., Yang, Y., Zhou, G. & Tu, Z. Immune suppression in chronic hepatitis B infection associated liver disease: A review. *World journal of gastroenterology* **25**, 3527-3537, doi:10.3748/wjg.v25.i27.3527 (2019).
- 38 Ji, Y. *et al.* Serum from Chronic Hepatitis B Patients Promotes Growth and Proliferation via the IGF-II/IGF-IR/MEK/ERK Signaling Pathway in Hepatocellular Carcinoma Cells. *Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology* **47**, 39-53, doi:10.1159/000489744 (2018).
- 39 Yoo, Y. *et al.* Hepatitis B virus X protein enhances transcriptional activity of hypoxia-inducible factor-1alpha through activation of mitogen-activated protein kinase pathway. *The Journal of biological chemistry* **278**, 39076-39084, doi:10.1074/jbc.M305101200 (2003).
- 40 Moon, E. *et al.* Hepatitis B virus X protein induces angiogenesis by stabilizing hypoxia-inducible factor-1alpha. *FASEB journal : official publication of the Federation of American Societies for Experimental Biology* **18**, 382-384, doi:10.1096/fj.03-0153fje (2004).

- 41 Chin, R. *et al.* Failure of lamivudine to reverse hepatitis B virus-associated changes in ERK, Akt and cell cycle regulatory proteins. *Antiviral therapy* **13**, 221-230 (2008).
- 42 Ru, J. *et al.* TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. *J Cheminform* **6**, 13, doi:10.1186/1758-2946-6-13 (2014).
- 43 Walters, J. R. New advances in the molecular and cellular biology of the small intestine. *Curr Opin Gastroenterol* **18**, 161-167, doi:10.1097/00001574-200203000-00002 (2002).
- 44 Xu, X. *et al.* A novel chemometric method for the prediction of human oral bioavailability. *Int J Mol Sci* **13**, 6964-6982, doi:10.3390/ijms13066964 (2012).
- 45 Kim, S. *et al.* PubChem Substance and Compound databases. *Nucleic Acids Res* **44**, D1202-1213, doi:10.1093/nar/gkv951 (2016).
- 46 O'Boyle, N. M. *et al.* Open Babel: An open chemical toolbox. *J Cheminform* **3**, 33, doi:10.1186/1758-2946-3-33 (2011).
- 47 Daina, A., Michielin, O. & Zoete, V. SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. *Nucleic Acids Res* **47**, W357-W364, doi:10.1093/nar/gkz382 (2019).
- 48 Szklarczyk, D. *et al.* STITCH 5: augmenting protein-chemical interaction networks with tissue and affinity data. *Nucleic Acids Res* **44**, D380-384, doi:10.1093/nar/gkv1277 (2016).
- 49 Stelzer, G. *et al.* The GeneCards Suite: From Gene Data Mining to Disease Genome Sequence Analyses. *Curr Protoc Bioinformatics* **54**, 1 30 31-31 30 33, doi:10.1002/cpbi.5 (2016).
- 50 Pinero, J. *et al.* The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res* **48**, D845-D855, doi:10.1093/nar/gkz1021 (2020).
- 51 Brown, G. R. *et al.* Gene: a gene-centered information resource at NCBI. *Nucleic Acids Res* **43**, D36-42, doi:10.1093/nar/gku1055 (2015).
- 52 UniProt, C. Reorganizing the protein space at the Universal Protein Resource (UniProt). *Nucleic Acids Res* **40**, D71-75, doi:10.1093/nar/gkr981 (2012).
- 53 Szklarczyk, D. *et al.* STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res* **47**, D607-D613, doi:10.1093/nar/gky1131 (2019).
- 54 Kohl, M., Wiese, S. & Warscheid, B. Cytoscape: software for visualization and analysis of biological networks. *Methods Mol Biol* **696**, 291-303, doi:10.1007/978-1-60761-987-1_18 (2011).

- 55 Zhang, Y. *et al.* A systems biology-based investigation into the therapeutic effects of Gansui Banxia Tang on reversing the imbalanced network of hepatocellular carcinoma. *Scientific reports* **4**, 4154, doi:10.1038/srep04154 (2014).
- 56 Xiao, Q., Wang, J., Peng, X., Wu, F. & Pan, Y. Identifying essential proteins from active PPI networks constructed with dynamic gene expression. *BMC genomics*, S1, doi:10.1186/1471-2164-16-s3-s1 (2015).
- 57 Saito, R. *et al.* A travel guide to Cytoscape plugins. *Nat Methods* **9**, 1069-1076, doi:10.1038/nmeth.2212 (2012).
- 58 Ashburner, M. *et al.* Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nat Genet* **25**, 25-29, doi:10.1038/75556 (2000).
- 59 Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res* **44**, D457-462, doi:10.1093/nar/gkv1070 (2016).
- 60 Dennis, G., Jr. *et al.* DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol* **4**, P3 (2003).
- 61 Berman, H., Henrick, K. & Nakamura, H. Announcing the worldwide Protein Data Bank. *Nat Struct Biol* **10**, 980, doi:10.1038/nsb1203-980 (2003).
- 62 Adams, P. D. *et al.* Announcing mandatory submission of PDBx/mmCIF format files for crystallographic depositions to the Protein Data Bank (PDB). *Acta Crystallogr D Struct Biol* **75**, 451-454, doi:10.1107/S2059798319004522 (2019).
- 63 Trott, O. & Olson, A. AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *Journal of computational chemistry* **31**, 455-461, doi:10.1002/jcc.21334 (2010).
- 64 Yuan, S., Chan, H. C. S., Filipek, S. & Vogel, H. PyMOL and Inkscape Bridge the Data and the Data Visualization. *Structure* **24**, 2041-2042, doi:10.1016/j.str.2016.11.012 (2016).
- 65 Barrett, T. *et al.* NCBI GEO: archive for functional genomics data sets—update. *Nucleic acids research* **41**, D991-995, doi:10.1093/nar/gks1193 (2013).
- 66 Zhou, W. *et al.* Predictive model for inflammation grades of chronic hepatitis B: Large-scale analysis of clinical parameters and gene expressions. *Liver Int* **37**, 1632-1641, doi:10.1111/liv.13427 (2017).

Figures

Figure 1

The flow diagram of network pharmacology analysis of TGYP-CHB related targets. TGYP target genes were firstly screened out according active compounds, and then performed with GO and KEGG pathways enrichment analysis. Next, CHB target genes were obtained from NCBI Gene, GeneCards and DisGeNET databased. The OEGs of TGYP and CHB target genes were isolated to identify key target genes. At last, the biologic process and expression of key target genes, interactions between of key compounds and key target genes were examined.

Figure 2

Functional Enrichment of TGYP Targets. Functional annotation of active compounds related to putative targets. (A) Enriched cellular components. (B) Enriched molecular functions. (C) Enriched biological processes. (D) Enriched KEGG pathways.

Figure 3

Identification of the Key Module. Process of identifying key modules. (A) Protein- protein interaction network. Blue hexagons represent weak interacting target genes, yellow hexagons represent strong interacting target genes; (B) Preliminary hub network. Yellow hexagons represent weak interacting target genes, red hexagons represent strong interacting target genes; (C) Hub network. Red circles represent related target genes, purple circles represent core target genes; (D) Key module. Purple circles represent key target genes.

Figure 4

Functional Enrichment of Key Targets. Functional annotation of key target genes. (A) Enriched cellular components. (B) Enriched molecular functions. (C) Enriched biological processes. (D) Enriched KEGG pathways.

Figure 5

The molecular docking of key compounds and key target genes. A. Interaction network of key target genes and corresponding active compounds of prediction. The green octagons represent key target genes, the yellow squares represent active compounds of prediction. The size of the icon was set according to the degree of the network. B. The structure of key compounds. C. The molecular docking of key compounds and key target genes.

Figure 6

Expression of key genes in CHB and correlation with liver inflammation. A. The upregulation of key target genes in CHB liver tissues. B. The key target genes expression was correlated to liver inflammation in CHB.

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

- [Supplementary.docx](#)
- [TableS1.xlsx](#)
- [TableS2.xlsx](#)
- [TableS3.xlsx](#)
- [TableS4.xlsx](#)