



OPEN

## Target identification of hepatic fibrosis using Pien Tze Huang based on mRNA and lncRNA

Jinhang Zhu<sup>1,2,4</sup>, Di Zhang<sup>1,2,4</sup>, Ting Wang<sup>2</sup>, Zhiliang Chen<sup>3</sup>, Luan Chen<sup>2</sup>, Hao Wu<sup>2</sup>, Cong Huai<sup>2</sup>, Jing Sun<sup>2</sup>, Na Zhang<sup>2</sup>, Muyun Wei<sup>2</sup>, Fei Hong<sup>3</sup>✉ & Shengying Qin<sup>1,2</sup>✉

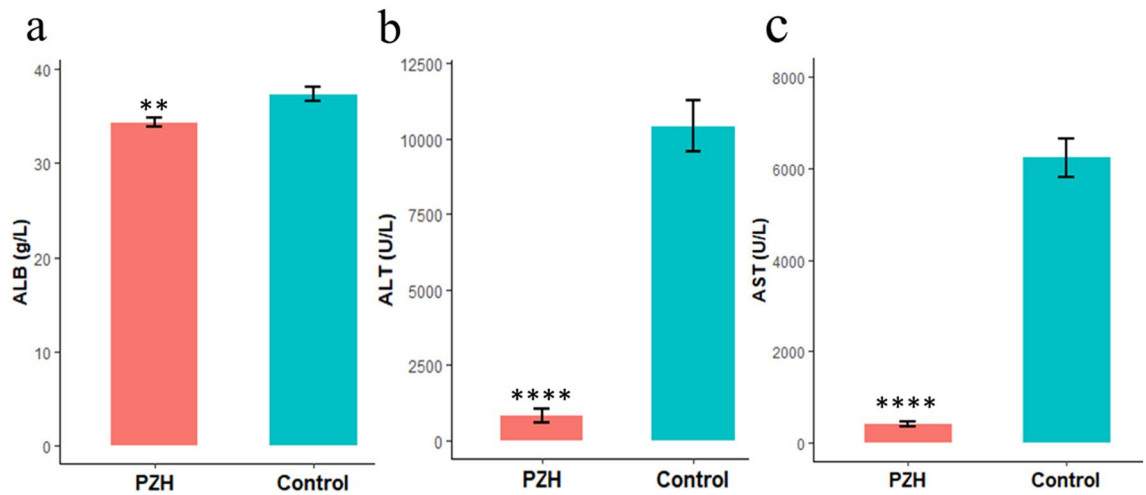
Hepatic fibrosis is a spontaneous wound-healing response triggered by chronic liver injury. Pien Tze Huang (PZH), a traditional Chinese herbal medicine, has been widely used to treat various hepatic diseases in Asia. We used a CCl<sub>4</sub>-induced mouse model to establish a PZH group of hepatic fibrosis mice treated with PZH and a control group of hepatic fibrosis mice without any treatment. We performed RNA-seq and mass spectrometry sequencing to investigate the mechanism of the PZH response in hepatic fibrosis and identified multiple differentially expressed transcripts (DETs) and proteins (DEPs) that may be drug targets of PZH. Liver functional indices, including serum albumin (ALB), alanine aminotransferase (ALT) and aspartate aminotransferase (AST), were significantly decreased in the PZH treatment group ( $P < 0.05$ ) in the eighth week. Hematoxylin–eosin (HE), Masson and Sirius red staining demonstrated that PZH significantly inhibited infiltration of inflammatory cells and collagen deposition. A total of 928 transcripts and 138 proteins were differentially expressed in PZH-treated mice compared to the control group. Gene Ontology (GO) enrichment analysis suggested that PZH may alleviate liver injury and fibrosis by enhancing the immune process. Taken together, our results revealed that multiple DETs and DEPs may serve as drug targets of PZH in hepatic fibrosis patient in future clinical practice.

Hepatic fibrosis is a spontaneous wound-healing response to multiple harmful stimuli, including long-term alcohol abuse, exposure to hepatotoxin, autoimmune response, and viral infection, and it has attracted increasing attention from clinicians and scientists<sup>1–3</sup>. Hepatic fibrosis will progress to liver cirrhosis and even hepatic carcinoma in the absence of an effective or timely treatment<sup>4,5</sup>. Animal models that simulate the causes of hepatic fibrosis have improved the understanding of liver injury, and CCl<sub>4</sub>-induced hepatic fibrosis is widely used in mouse models to study hepatotoxic mechanisms<sup>6,7</sup>.

Currently, treatments for hepatic fibrosis mainly include eliminating the primary causes and suppressing inflammation. However, clinical treatment for hepatic fibrosis lacks appropriate methods<sup>8,9</sup>. Pien Tze Huang (PZH), a well-known herbal traditional Chinese medicine, is widely used in Asia for various liver diseases, including hepatic fibrosis, cirrhosis, hepatocellular carcinoma, and liver failure<sup>10,11</sup>. Anti-inflammation is an important therapeutic function of PZH drugs<sup>12</sup>. PZH has been widely utilized as a therapy for liver diseases for over 600 years, and it primarily contains musk, gallstone of the ox, San qi and snake's gall<sup>13</sup>. However, the underlying mechanism of how PZH regulates gene expression in hepatic fibrosis and the related biomarkers remains unclear.

Several studies in vitro and in animal models have demonstrated that PZH could alleviate various liver diseases. For example, Jinyan Zhao et al. found that PZH ameliorates CCl<sub>4</sub>-induced liver injury, oxidative stress and inflammation in rats<sup>12</sup>; Yang Yang et al. found that PZH could ameliorate liver injury by inhibiting the PERK/eIF2 $\alpha$  signaling pathway in alcohol and high-fat diet rats<sup>13</sup>; and Haiyin Zheng et al. found that PZH ameliorates hepatic fibrosis by suppressing the NF- $\kappa$ B pathway using real-time PCR experiment methods based on rat and cell lines<sup>14</sup>. However, PCR and other traditional experimental methods cannot be used to identify treatment biomarkers of hepatic fibrosis at the whole-transcriptome level. Therefore, high-throughput whole transcriptome sequencing (RNA-seq) is urgently needed for more biomarker identification.

<sup>1</sup>Department of Physiology, School of Basic Medical Sciences, Anhui Medical University, Hefei 230032, China. <sup>2</sup>Bio-X Institutes, Key Laboratory for the Genetics of Developmental and Neuropsychiatric Disorders, Ministry of Education, Shanghai Jiao Tong University, Shanghai, China. <sup>3</sup>Fujian Provincial Key Laboratory of PTH Natural Medicine Research and Development, Zhangzhou PTH Pharmaceutical CO., LTD, Zhangzhou, China. <sup>4</sup>These authors contributed equally: Jinhang Zhu and Di Zhang. ✉email: pzhhf123@126.com; chinsir@situ.edu.cn



**Figure 1.** Liver function serum biochemical indices at the time point of eighth week. (a) Serum ALB indices of the PZH and control groups. (b) Serum ALT indices of the PZH and control groups. (c) Serum AST indices of the PZH and control groups. Mean  $\pm$  SE. \*\* $P < 0.01$ , \*\*\*\* $P < 1.0 \times 10^{-4}$ ,  $n = 6$ .

RNA-seq has become a reliable option to explore drug effects based on transcriptome changes. The technology has been widely used to identify specific pharmacological effects<sup>15,16</sup>. Messenger RNAs (mRNAs), which are transcribed by corresponding protein-coding genes, influence the different kinds of drug-metabolizing enzymes and the translation efficiency of genes<sup>17</sup>. Long noncoding RNAs (lncRNAs) are regarded as regulators that participate in disease development and biological functions<sup>18,19</sup>. Analysis of mRNA and lncRNA expression is now an important method to identify or explore their regulatory roles in the mechanisms of various diseases<sup>20–22</sup>.

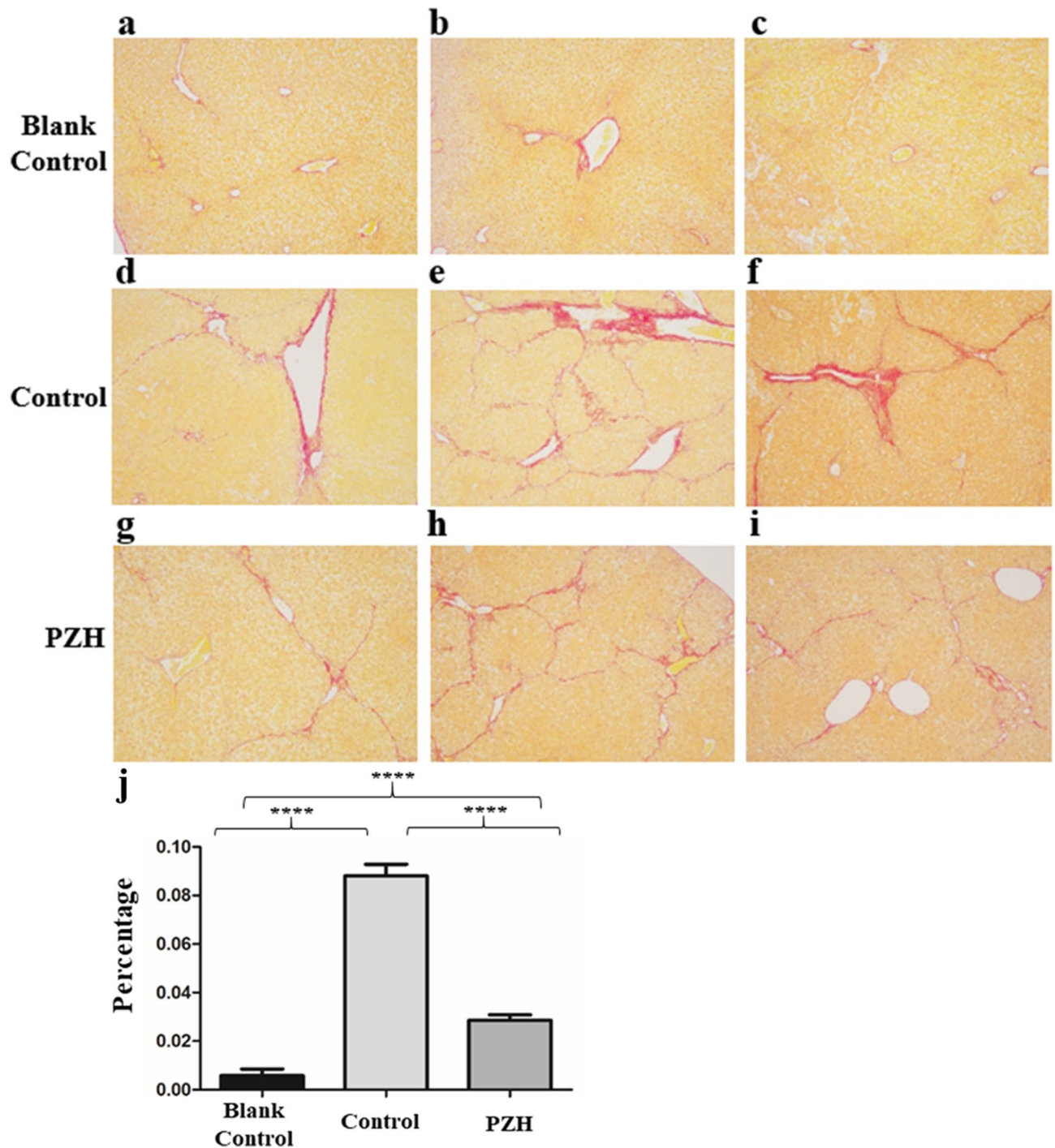
Here, we performed a case–control study using a CCl<sub>4</sub>-induced hepatic fibrosis mouse model to explore the drug targets and pathways of PZH. After transcriptome sequencing and mass spectrometry sequencing data analysis, 928 DETs and 138 DEPs were identified. Through functional annotation enrichment analysis, immune system biological process was significantly enriched based on both DETs of RNA-seq and DEPs of mass spectrometry sequencing. These results revealed that PZH alleviated hepatic injury and fibrosis by improving the immune system biological processes. Our study provides potential transcriptional therapeutic targets for the clinical application of hepatic fibrosis.

## Results

**Liver functional indices were significantly decreased in the PZH group compared to the control group.** The liver function indices of 3 mice on the second, fourth and sixth weeks are shown in Supplementary Table 1. Their  $P$  values were not significantly changed between the PZH and control groups ( $P > 0.05$ ). Differing from the ones measured in the second, fourth and sixth week, the liver function indexes of 6 mice of the eighth week were significantly increased in control groups (ALB:  $P = 9.76 \times 10^{-3}$ , ALT:  $P = 4.43 \times 10^{-5}$ , AST:  $P = 2.45 \times 10^{-5}$ , Fig. 1, Supplementary Table 2). These results mean that the level of CCl<sub>4</sub>-induced liver injury in the control group was significantly higher than that in the PZH group at the eighth week.

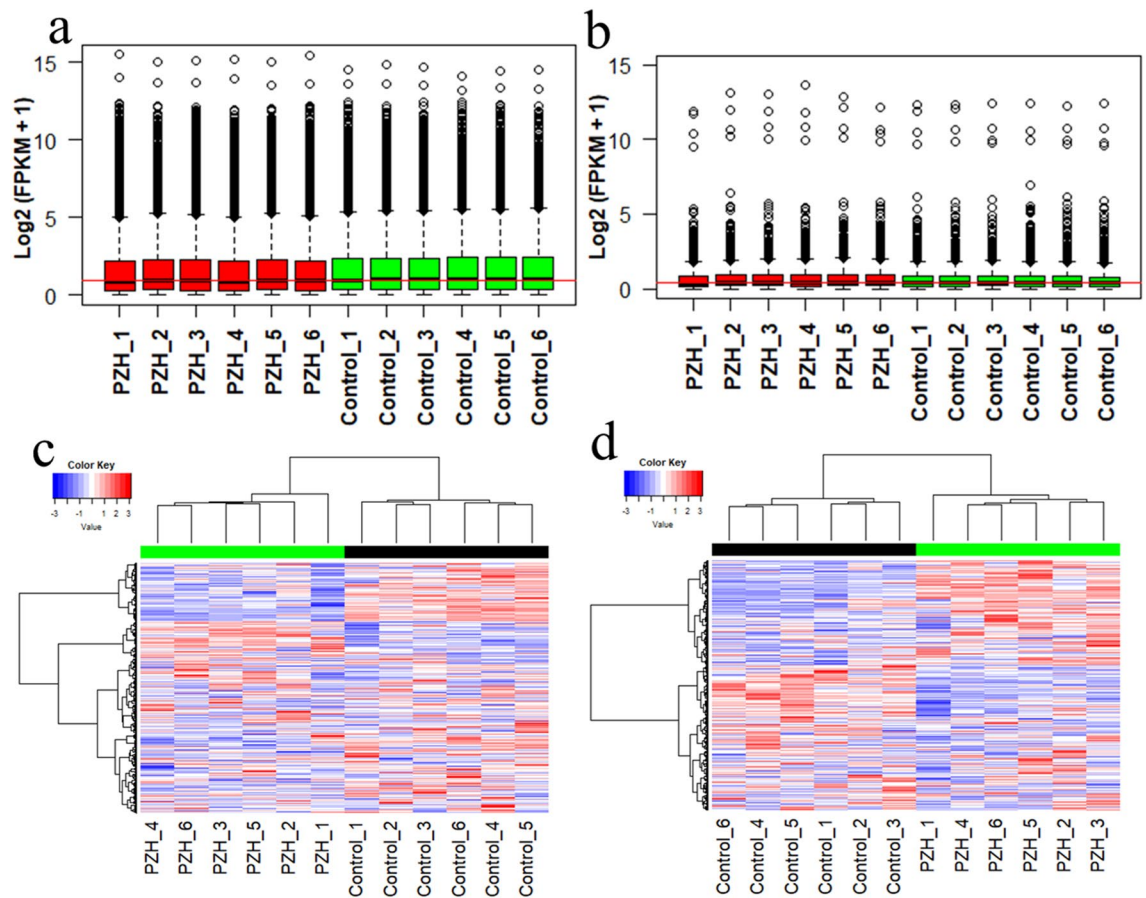
**Comparing HE, Masson and Sirius red staining results of the PZH and control groups.** HE and Masson staining results of the PZH and control groups at the second, fourth, sixth and eighth weeks are shown in the Supplementary Information (HE: Supplementary Figs. 1–4; Masson: Supplementary Figs. 5–8,  $N = 3$ ). These results demonstrated that PZH treatment gradually alleviated the progression of hepatic fibrosis compared to the process observed in the control group. Based on HE staining results from 3 mice of the eighth week after treatment, granules with less blue color appeared around the central veins of the liver in the hepatic fibrosis mice of the PZH group than those of the control group. This result demonstrated that infiltrating inflammatory cells were less abundant in the liver tissue of the PZH group compared with the control group (Supplementary Fig. 4). The Masson staining results demonstrated that the hepatic tissue of 3 mice in the PZH group appeared to have fewer blue zones around the central veins of the liver compared to those in the control group (Supplementary Fig. 8). Besides, we performed the Sirius red staining and the results showed that the ratio of positive area of Sirius red staining accounting for total area was significantly decreased in PZH treatment group compared with non-treatment group (Fig. 2). Taken together, the HE, Masson and Sirius red staining results revealed that PZH significantly inhibited infiltration by inflammatory cells and collagen deposition in the CCl<sub>4</sub>-induced mouse liver fibrosis model.

**Boxplots and hierarchical cluster analysis of the overall expression of protein-coding and lncRNA transcripts.** To identify related biomarkers, RNA-seq was performed based on the liver samples of 12 mice including 6 mice treatment using PZH and 6 mice without any treatment of CCl<sub>4</sub>-induced mice model at the time point of eighth week. The overall expression levels of protein-coding and lncRNA transcripts were visualized by boxplots and hierarchical clustering analysis. The results revealed that the overall expression levels



**Figure 2.** Sirius red staining results of hepatic fibrosis mice in the PZH and control groups at the time point of eighth week. (a–c) Three Sirius red staining results of the blank control group which received neither  $\text{CCl}_4$  nor PZH. (d–f) Three Sirius red staining results of the control group which received  $\text{CCl}_4$ . (g–i) Three Sirius red staining results of the PZH group which received both  $\text{CCl}_4$  and PZH. (j) Was the percentage of positive area of Sirius red staining accounting for total area of blank control, control and PZH group (\*\*\*\* $P < 1.0 \times 10^{-4}$ ).

of protein-coding transcripts were higher than the overall expression levels of lncRNA transcripts, and the overall expression levels of protein-coding and lncRNA transcripts did not significantly change between the PZH and control groups (Fig. 3a,b). The six samples from the PZH group of eighth week were clustered together using hierarchical cluster analysis. Similarly, the six samples from the control group of eighth week were clustered together. The results also suggested that the expression profiles of the biological replicates of each group were highly consistent (Fig. 3c,d), and the drug effect was the main cause of the differentially expressed transcripts between the PZH and control groups.



**Figure 3.** Visualization of the overall expression levels across all the samples at the time point of eighth week. (a) Expression levels of protein-coding transcripts and (b) expression levels of lncRNA transcripts; the x-axis indicates the names of the 12 samples, and the y-axis indicates the  $\log_2(\text{FPKM} + 1)$  value. (c) Hierarchical clustering result of the protein-coding transcripts. (d) Hierarchical cluster result of lncRNA transcripts. The mean was set to 0, and the standard deviation was set to 1 in the calculation of the  $\log_2(\text{FPKM} + 1)$  value of each gene in the 12 samples. Red indicates a high expression level, while blue represents a low expression level. Heat maps of (c) and (d) were created using R programming language (R: A Language and Environment for Statistical Computing, R Core Team, R Foundation for Statistical Computing, Vienna, Austria, 2016, <https://www.R-project.org>).

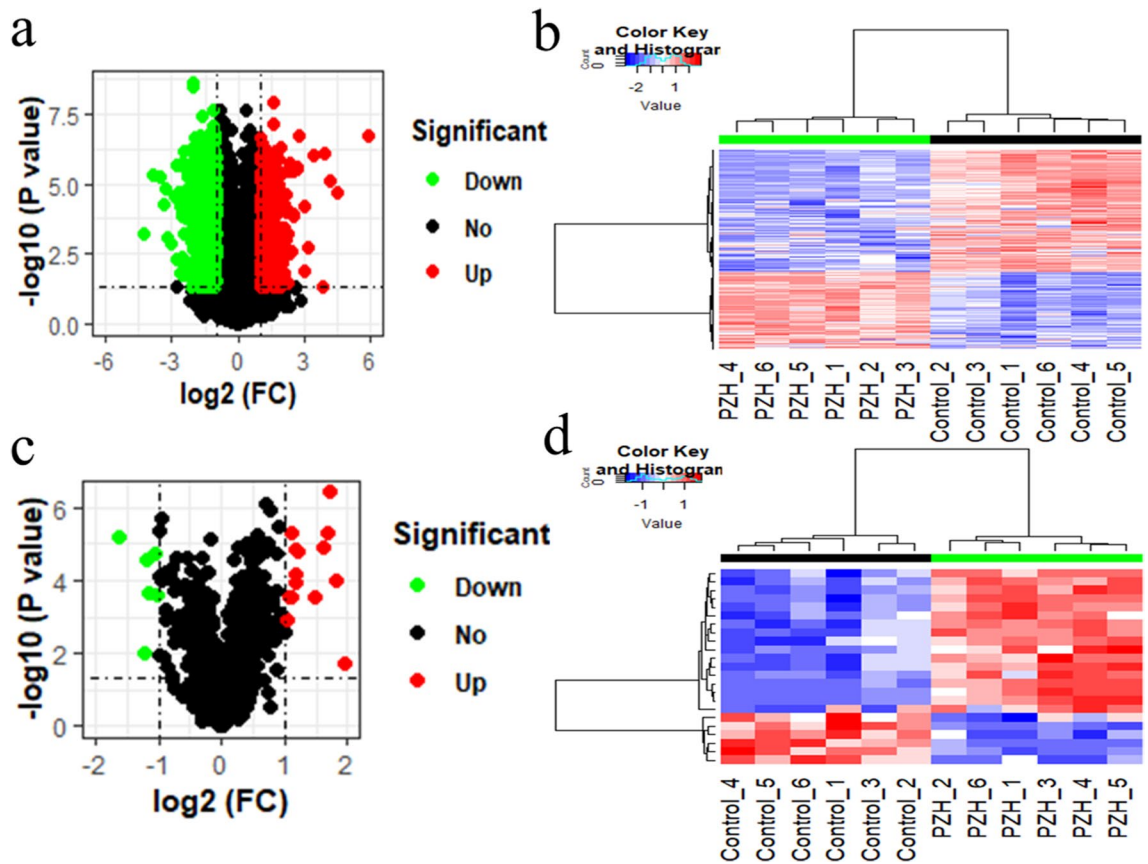
### Identification of differentially expressed protein-coding and lncRNA transcripts between the PZH and control groups.

We identified the differentially expressed protein-coding and lncRNA transcripts between 6 PZH with 6 control liver tissues at the time point of eighth week. The number of differentially expressed protein-coding and lncRNA transcripts was visualized using volcano plots and heat maps. A differentially expressed transcript was defined as one with fold change (FC)  $> 2$  or  $\text{FC} < -2$ , and the  $P$  value threshold was set to 0.05. A total of 558 significantly upregulated and 347 significantly downregulated protein-coding transcripts were successfully identified (Fig. 4a and Supplementary Table 3). Their expression levels are shown in Fig. 4b. In addition, 6 significantly upregulated and 17 downregulated lncRNA transcripts were identified (Fig. 4c and Supplementary Table 4), and their expression levels are shown in Fig. 4d.

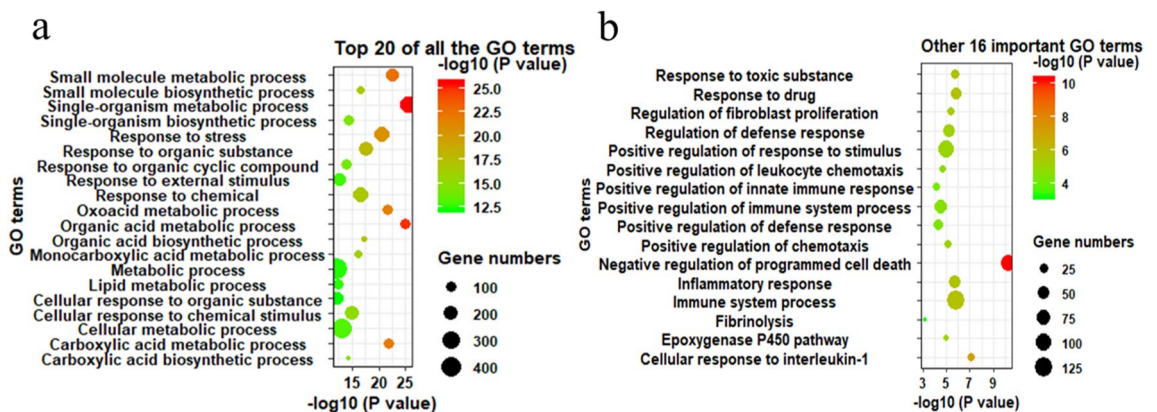
**Functional annotation of the differentially expressed protein-coding transcripts.** The results of the GO enrichment analysis were visualized using bubble plots, and the top 20 GO terms are shown in Fig. 5a and Supplementary Table 5. It appears that the most enriched annotated pathways were predominately related to metabolism. Besides, positive regulation of the immune system process, positive regulation of the defense response, response to toxic substances, fibrinolysis and fibroblast proliferation biological processes were enriched, and these GO terms explained the potential therapeutic mechanism of PZH in hepatic fibrosis (Fig. 5b and Supplementary Table 6).

**Co-expression network construction with differentially expressed lncRNAs and protein-coding transcripts.** We constructed a co-expression network using 23 differentially expressed lncRNAs and 905 differentially expressed protein-coding transcripts. Three co-expression modules were identified as shown Fig. 6a. These three co-expression modules included a module that had 242 transcripts (marked in blue), module that had 135 transcripts (marked in brown color) and module that had 551 transcripts (marked in turquoise).





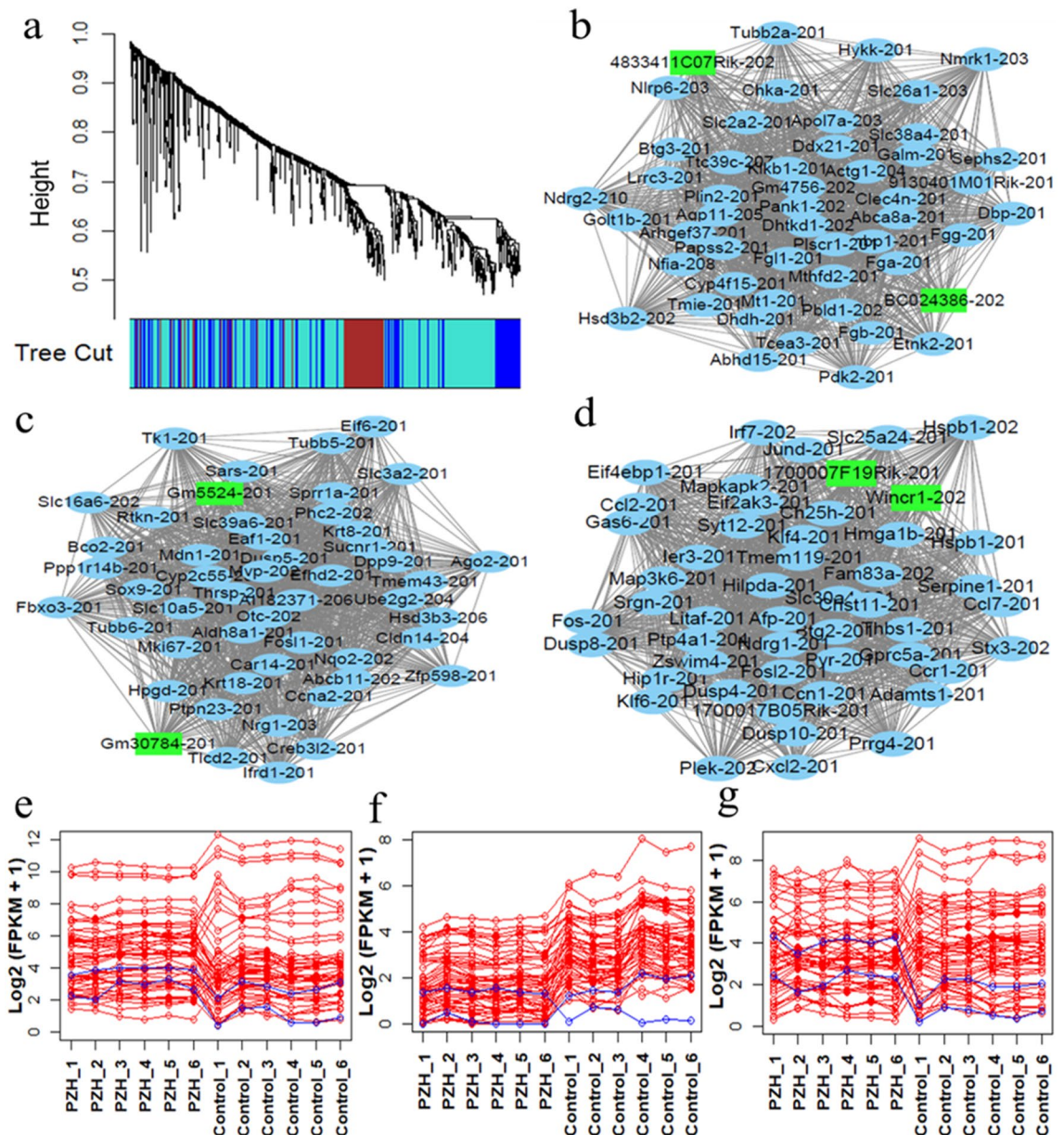
**Figure 4.** Visualization of the differentially expressed transcripts. (a) Volcano plot displays the differentially expressed protein-coding transcripts. (b) Heat map displays the expression level of differentially expressed protein-coding transcripts. (c) Volcano plot displays the differentially expressed lncRNA transcripts. (d) Heat map displays the expression of differentially expressed lncRNA transcripts. The red dots represent upregulated transcripts while green dots indicate downregulated transcripts in volcano plot. Red indicates high expression level while blue represents low expression level in the heat map. Heat maps of (b) and (d) were created using R programming language (R: A Language and Environment for Statistical Computing, R Core Team, R Foundation for Statistical Computing, Vienna, Austria, 2016, <https://www.R-project.org>).



**Figure 5.** Visualization of the functional annotation of differentially expressed protein-coding transcripts. (a) The top 20 significantly enriched GO biological process terms. (b) The other 16 important enriched GO biological process terms.

color). Each co-expression module represented a group of transcripts with similar expression profiles across all the samples.

The top 50 hub nodes were extracted from each module based on analysis of the co-expression network between the lncRNAs and protein-coding DETs of each module. We constructed a network of each module using Cytoscape software (Fig. 6b–d). The network results revealed that the differentially expressed lncRNA transcripts

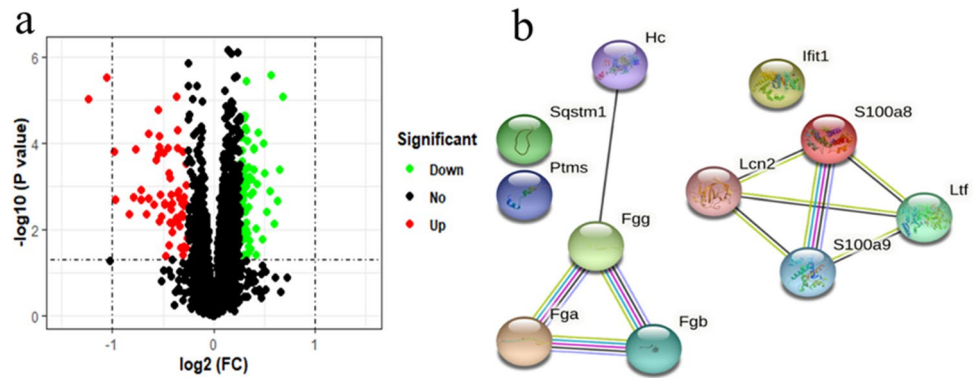


**Figure 6.** Co-expression network of the differentially expressed lncRNA and protein-coding transcripts. (a) Visualization of the co-expression modules of all the differentially expressed lncRNA and protein-coding transcripts; each color represents one module. (b–d) Visualization of the top 50 hub transcripts in the blue, brown and turquoise modules. The green nodes represent lncRNA transcripts, and the blue nodes represent protein-coding transcripts. (e–g) Expression patterns of the top 50 hub transcripts in each module. The x-axis indicates the names of each sample, and the y-axis indicates the expression  $\log_2(\text{FPKM} + 1)$  value. The blue lines represent differentially expressed lncRNA transcripts, and the green lines represent differentially expressed protein-coding transcripts.

interacted with the differentially expressed protein-coding transcripts. The expression pattern of each of the top 50 transcripts of the three modules is shown in Fig. 6e–g. The expression pattern revealed that these differentially expressed lncRNA transcripts were co-expressed with the differentially expressed protein-coding transcripts. The absolute correlation values of the top 50 hub transcripts of the blue, brown and turquoise modules were 0.91 ~ 1.0, 0.90 ~ 1.0 and 0.93 ~ 1.0, respectively.

**Differentially expressed proteins (DEPs) involve in immune system process.** We detected the protein expression using mass spectrometry sequencing based on the liver tissues of 6 treatment mice using PZH and 6 mice without any treatment of CCl<sub>4</sub>-induced mice model at the time point of eighth week, and 138 DEPs were identified (Fig. 7a and Supplementary Table 7). Function enrichment analysis showed that immune system biological process were significantly enriched based on the 138 DEPs ( $P = 3.52 \times 10^{-4}$ ), and this biological process was also significantly enriched based on DEGs of RNA-seq. Plasminogen activation and fibrinolysis biology pro-





**Figure 7.** Visualization of the differentially expressed proteins. (a) Volcano plot displays the differentially expressed protein-coding proteins. (b) Protein–protein interaction network of the 11 genes which involved in the immune system, this network were constructed using STRING database which is publicly available (STRING version: 11.0<sup>23</sup>).

cess also were significant enriched. Protein reflect the final biological function of gene, therefore, we believe that PZH treatment liver fibrosis through enhance the immune system. Protein–protein interaction network were constructed based on the 11 genes which involved in the immune system (Fig. 7b).

## Discussion

Although PZH has been widely used to treat liver disease in Asia for hundreds of years, its transcriptional drug targets are still unknown<sup>14,24</sup>. In our study, we designated two groups of hepatic fibrosis mice established using CCl<sub>4</sub>. One group received PZH treatment, while the other group did not. We found that liver function indices were obviously decreased in the PZH group, and the results suggested that PZH treatment significantly attenuated liver injury. HE staining results revealed that severe inflammation occurred in the control group. Masson staining also showed that a large number of blue collagenous fibers were deposited surrounding the hepatic vein in the control group. Sirius red staining results showed that the ratio of positive area of Sirius red staining accounting for total area was significantly decreased in PZH treatment group compared with non-treatment group. These results demonstrated that PZH treatment could relieve CCl<sub>4</sub>-induced liver injury, inflammation and hepatic fibrosis.

To identify the drug targets by transcriptome and protein level analysis, we first performed RNA-seq and mass spectrometry sequencing based on the liver tissue of the PZH and control groups. We found that 928 transcripts and 138 proteins were differentially expressed. Functional enrichment analysis showed that positive regulation of immune system process were significantly enriched based on both 905 DETs and 138 DEPs. This result revealed that PZH could alleviate hepatic fibrosis by enhancing the immune system biological processes. Previous studies have reported the hypothesis that PZH might enhance the immune system, which is considered to be the functional mechanism of PZH<sup>13,25</sup>, and we confirmed this hypothesis and identified the exact DETs in these biological processes. Besides, plasminogen activation and fibrinolysis biological process were also significantly enriched based on the 138 DEPs, these two biological processes may be also participate in the treatment process of hepatic fibrosis using PZH.

Co-expression means that the genes interact with each other<sup>26</sup>. Because of the lack of an integrated lncRNA function annotation database, we used the WGCNA package to find differentially expressed lncRNAs co-expressed with protein-coding transcripts<sup>19,27</sup>. Three modules were constructed based on these 23 differentially expressed lncRNAs and 905 protein-coding transcripts, and the top 50 hub transcripts of each module were extracted. Analysis of the top 50 hub transcripts of each module revealed that 6 differentially expressed lncRNA transcripts were highly correlated with the differentially expressed protein-coding transcripts of each pattern. These 6 differentially expressed lncRNA transcripts were reported as PZH therapeutic drug targets for the first time in this study.

As a high-throughput method, RNA-seq and mass spectrometry sequencing technology has helped researchers find increasingly more drug targets. In our study, we identified 928 DETs and 138 DEPs, and we hypothesized that these differentially expressed transcripts and proteins were either the drug targets or the injury-dependent effects of PZH ameliorating hepatic fibrosis. A limitation of this study is that more molecular biology experiments are necessary to validate these results. However, we cannot validate so many DETs or DEPs one by one. To overcome this limitation, we researched related papers and found that multiple DETs or DEPs identified in our study have already been found to participate in hepatic fibrogenesis using molecular biological experiment methods in previous studies. For example, Ximeng Han et al. reported that SOX9<sup>+</sup> hepatocytes have repair and regeneration functions after liver injury<sup>28</sup>. Inagaki et al. revealed the mechanisms of the Smad2 gene in hepatic fibrogenesis<sup>29</sup>. Wu et al. reported that quercetin prevented CCl<sub>4</sub>-induced hepatic fibrosis by regulating TIMP1 gene expression<sup>30</sup>. The genes ADAMTS1, Cyp2c55, Serpinh1, MMP subfamily, Gpt, IL1B, and THBD have also been reported by previous studies<sup>31–37</sup>. These results further validated our findings. Nevertheless, the identified DETs in our study provide many new potential drug targets for scientists to further investigate.

In summary, we constructed two groups of CCl<sub>4</sub>-induced hepatic fibrosis mouse models: one group was treated with PZH and the other group was not treated with PZH. Finally, we found that the control group suffered obvious hepatic fibrosis, while the PZH group had very little hepatic fibrosis in the eighth week. We identified 905 differentially expressed protein-coding transcripts, 23 differentially expressed lncRNA transcripts and 138 DEPs. These transcripts and DEPs may be the drug targets of PZH ameliorating liver injury and fibrosis. We demonstrated that PZH positively regulates the immune system biological process based on both DETs of RNA-seq and DEPs of mass spectrometry sequencing. Our study provides potential transcriptional and protein level therapeutic targets to guide the clinical application of PZH.

## Methods

**Preparation of reagents.** PZH was obtained from and authenticated by Zhangzhou Pien Tze Huang Pharmaceutical Co., Ltd. (Zhangzhou, China; Chinese Food and Drug Administration approval no. Z35020242). Other required reagents, including staining reagents and antibodies, were purchased from other companies (Cell Signaling Technology; Wuhan Yafa Biological Technology Co., Ltd.). Thirty C57BL/6J male mice were obtained from laboratory of Shanghai Southern Model Animal Center, and they were SPF level and six weeks old. The PZH drug powder was dissolved in double distilled water, and its concentration was diluted to 25 mg/ml. Carbon tetrachloride (CCl<sub>4</sub>) was dissolved in 10% corn oil (1:1 vol/vol). These prepared reagents were stored at -80 °C until further usage. The doses of PZH and the mixture of CCl<sub>4</sub> with corn oil were 0.25 mg/g and 10 µl/g, respectively, according to the weight of each mouse.

**Randomized controlled mouse hepatic fibrosis model using CCl<sub>4</sub> and PZH plus CCl<sub>4</sub>.** Thirty mice were randomly divided into two groups named the ‘PZH group’ and ‘control group’, respectively. All of these 30 mice were 6 weeks C57BL male mice. The 15 mice in the control group received intraperitoneal injection of 10 µl/g solvent mixture of CCl<sub>4</sub> and 10% corn oil twice a week and intragastric administration of double distilled water once a day. The 15 mice in the PZH group received intraperitoneal injection of 10 µl/g solvent mixture of CCl<sub>4</sub> and 10% corn oil twice a week and intragastric administration of 0.25 mg/g of the PZH prepared reagent once a day. During this 8-week experimental period, 3 mice in each group were sacrificed by the cervical dislocation method every two weeks. After 8 weeks, the remaining 12 mice were sacrificed by the cervical dislocation method, and their liver tissue and blood were harvested. The detailed information of mice body weight was showed in Supplementary Table 8, and the body weight result showed that there no significant adverse effect in PZH treatment group.

**Hepatic biochemical index detection.** The levels of hepatic biochemical indices, aspartate aminotransferase (AST), alanine aminotransferase (ALT) and albumin (ALB), were determined by ELISA kits (Roche; Shanghai). These indices were detected every two weeks to estimate the extent of liver injury. AST and ALT mainly exist in hepatic cells. Their levels were detected in blood when more hepatic cells were damaged. ALB was detected by the bromocresol green method. Sample blood was used to examine hepatic biochemical indices within 2 h after collection to ensure the reliability of the results by the Roche Cobas 8000 automatic biochemical analyzer. The liver function indices measured in the second, fourth and sixth weeks are shown in Supplementary Table 1, and the liver function index results from the eighth week are shown in Supplementary Table 2.

**HE, Masson and Sirius red staining.** Pathological staining techniques have been regarded as the gold standard for diagnosing hepatic injury or fibrosis. To follow disease development and ensure that the liver fibrosis mouse model was constructed successfully, we used hematoxylin–eosin staining named ‘HE’, Masson’s trichrome staining called ‘Masson’ and Sirius red staining to validate the results of the mouse model. Liver tissue fragments of approximately 1 cm<sup>3</sup> extracted from mouse livers were fixed using 10% formalin. The liver tissues were embedded in paraffin wax (GeLin Company; Thermo Shandon HistoCentre 2), and the tissues were cut into 5 mm thick sections with a microtome (Thermo Fisher Scientific, Microm HM3550). After the tissue sections were dehydrated by increasing alcohol concentrations, the tissue sections were cleared using xylene (GeLin Company). The tissue sections were stained by HE, Masson and Sirius red staining. The cell nuclei were stained blue, and the cytoplasm was stained red or pink using HE staining. The marked colors were used to determine whether inflammatory cells had infiltrated into the hepatic central small vein. Masson staining was performed to identify various types of fibers. Different types of fibers were specifically stained in different colors by Masson staining in muscle tissue. Sections, made with paraffin-wax embedded liver tissue, were stained in picric acid Sirius red solution for 8 min, and rinse for several minutes with anhydrous alcohol. And then, tissue sections were placed in 60 oven, using xylene transparent 5 min and neutral gum seal. The extent of staining results for Sirius red was quantified using ImageJ software<sup>38</sup>.

**RNA extraction and sequencing.** To identify related biomarkers and therapeutic mechanisms, 12 mice were sacrificed on the eighth week by the cervical dislocation method to obtain their liver tissues with hepatic fibrosis. Total RNA from each hepatic fibrosis mouse tissue sample was extracted using TRIzol reagent (Invitrogen, CA, USA) following the manufacturer’s instructions. The integrity of total RNA was evaluated by the Bioanalyzer 2100 system (Agilent, CA, USA), and the quality scores are shown in Supplementary Fig. 9. The RNA integrity number (RIN) and 28S:18S ratio were first detected to ensure good RNA quality of each tissue to conduct the following sequencing analysis. All the RINs were larger than 9.1 (Supplementary Fig. 9a), and the 28S:18S ratios were larger than 1.2 (Supplementary Fig. 9b). The RNA library was constructed using strand-specific sequencing technologies following the manufacturer’s recommended protocol of the TruSeq RNA Sam-



ple Prep Kit. The RNA libraries were sequenced using the Illumina HiSeq 4000 platform, and 150 bp paired-end reads were generated.

**Quality control, mapping and transcript assembly based on trimmed high-quality RNA-seq data.** FastQC software was used to perform quality detection (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Clean high-quality data were obtained by removing adaptors, poly-N sequences, low-quality reads and fragments from raw data using Trimmomatic software<sup>39</sup>. Read quality is shown in Supplementary Fig. 10a. We then mapped the processed RNA-seq data to the Ensembl mouse reference genome (Mus\_musculus.GRCm38) using HiSat2 version 2.1.0 software<sup>40</sup>. The mapping ratio of each sample was larger than 93.2% (Supplementary Fig. 10b). Transcript assembly was performed by the Ensembl mouse annotation database (Mus\_musculus.GRCm38.95.gtf) using StringTie version 1.3.5 software<sup>41</sup>. Finally, the fragments per kilobase of transcript per million mapped reads (FPKM) expression matrix was obtained. A total of 29,883 protein-coding and 1432 lncRNA transcripts were annotated.  $\text{Log}_2(\text{FPKM} + 1)$  was used to normalize the expression data.

**Differentially expressed mRNA and lncRNA transcript identification and functional annotation.** Differentially expressed transcripts were identified using R. The genes with  $\text{FC} > 2$  or  $\text{FC} < -2$  and a  $P$  value  $< 0.05$  were defined to be significantly different between the two groups. The DAVID database was used to perform Gene Ontology (GO) enrichment analysis to annotate the functions of the differentially expressed transcripts (<https://david.ncicrf.gov/>). Currently, there is a lack of integrated functional annotation databases that could be used to annotate mouse lncRNA functions. Weighted gene co-expression network analysis (WGCNA), an R package, was used to explore the co-expression relationship between lncRNAs and protein-coding transcripts<sup>27,42,43</sup>. We constructed a network of each module using Cytoscape software<sup>44</sup>. The soft threshold of the WGCNA package was  $\beta = 14$ , as shown in Supplementary Fig. 11.

**Ethics approval.** The animal experiments were conducted in accordance with the guidelines of the National Institutes of Health of China for the care and use of laboratory animals, and all the animal work was approved and conducted under the guidelines of the Institutional Animal Care and Use Committee, Shanghai Jiao Tong University, China. We confirm that the study is reported in accordance with ARRIVE guidelines.

### Data availability

The RNA-seq dataset of 12 samples has been deposited in NCBI Gene Expression Omnibus (GEO) with accession code GSE133481.

Received: 2 March 2021; Accepted: 26 July 2021

Published online: 20 August 2021

### References

- Maepa, M. B., Ely, A. & Arbuthnot, P. How successful has targeted RNA interference for hepatic fibrosis been?. *Expert. Opin. Biol. Ther.* **18**, 381–388 (2018).
- Li, D. *et al.* Anti-fibrotic role and mechanism of *Periplaneta americana* extracts in CCl<sub>4</sub>-induced hepatic fibrosis in rats. *Acta Biochim. Biophys. Sin. (Shanghai)*. **50**, 491–498 (2018).
- Tao, R. *et al.* MicroRNA-29b-3p prevents *Schistosoma japonicum*-induced liver fibrosis by targeting COL1A1 and COL3A1. *J. Cell Biochem.* **119**, 3199–3209 (2018).
- Dongiovanni, P., Romeo, S. & Valenti, L. Hepatocellular carcinoma in nonalcoholic fatty liver: Role of environmental and genetic factors. *World J. Gastroenterol.* **20**, 12945–12955 (2014).
- Namisaki, T. *et al.* Liver fibrosis progression predicts survival in patients with primary biliary cirrhosis. *Hepatol. Res.* **47**, E178–E186 (2017).
- Shrestha, N. *et al.* Glutamine inhibits CCl<sub>4</sub> induced liver fibrosis in mice and TGF- $\beta$ 1 mediated epithelial–mesenchymal transition in mouse hepatocytes. *Food Chem. Toxicol.* **93**, 129–137 (2016).
- Mansour, M. F., Greish, S. M., El-Serafi, A. T., Abdelall, H. & El-Wazir, Y. M. Therapeutic potential of human umbilical cord derived mesenchymal stem cells on rat model of liver fibrosis. *Am. J. Stem Cells.* **8**, 7–18 (2019).
- Altamirano-Barrera, A., Barranco-Fragoso, B. & Mendez-Sanchez, N. Management strategies for liver fibrosis. *Ann. Hepatol.* **16**, 48–56 (2017).
- Bozic, M. & Molleston, J. Strategies for management of pediatric cystic fibrosis liver disease. *Clin. Liver Dis. (Hoboken)*. **2**, 204–206 (2013).
- Li, S. X., Mu, Y. & Zheng, F. Y. Influence of gastrointestinal digestion and edible plant combination on oral bioavailability of triterpene saponins, using a biomimetic digestion and absorption system and determination by HPLC. *J. Agric. Food Chem.* **61**, 10599–10603 (2013).
- Qi, F. *et al.* Pien Tze Huang inhibits the growth of hepatocellular carcinoma cells by upregulating miR-16 expression. *Oncol. Lett.* **14**, 8132–8137 (2017).
- Zhao, J. *et al.* Pien Tze Huang Gan Bao ameliorates carbon tetrachloride-induced hepatic injury, oxidative stress and inflammation in rats. *Exp. Ther. Med.* **13**, 1820–1826 (2017).
- Yang, Y. *et al.* Pien Tze Huang ameliorates liver injury by inhibiting the PERK/eIF2 $\alpha$  signaling pathway in alcohol and high-fat diet rats. *Acta Histochem.* **120**, 578–585 (2018).
- Zheng, H. *et al.* Pien-Tze-Huang ameliorates hepatic fibrosis via suppressing NF- $\kappa$ B pathway and promoting HSC apoptosis. *J. Ethnopharmacol.* **244**, 111856 (2019).
- Ye, C. *et al.* DRUG-seq for miniaturized high-throughput transcriptome profiling in drug discovery. *Nat. Commun.* **9**, 4307 (2018).
- Boswell, S. A. *et al.* Total RNA-seq to identify pharmacological effects on specific stages of mRNA synthesis. *Nat. Chem. Biol.* **13**, 501–507 (2017).
- Leppke, K., Das, R. & Barna, M. Functional 5' UTR mRNA structures in eukaryotic translation regulation and how to find them. *Nat. Rev. Mol. Cell Biol.* **19**, 158–174 (2018).
- Lee, J. T. Epigenetic regulation by long noncoding RNAs. *Science* **338**, 1435–1439 (2012).
- Xu, J. *et al.* A comprehensive overview of lncRNA annotation resources. *Brief Bioinform.* **18**, 236–249 (2017).

20. Grelet, S. *et al.* A regulated PNUTS mRNA to lncRNA splice switch mediates EMT and tumour progression. *Nat. Cell Biol.* **19**, 1105 (2017).
21. Wang, W. *et al.* Integrative analysis of mRNA and lncRNA profiles identified pathogenetic lncRNAs in esophageal squamous cell carcinoma. *Gene* **661**, 169–175 (2018).
22. Zhang, J., Le, T. D., Liu, L. & Li, J. Inferring and analyzing module-specific lncRNA-mRNA causal regulatory networks in human cancer. *Brief Bioinform.* **20**, 1403–1419 (2018).
23. 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 (2019).
24. Lin, W. *et al.* Pien Tze Huang inhibits liver metastasis by targeting TGF-beta signaling in an orthotopic model of colorectal cancer. *Oncol. Rep.* **33**, 1922–1928 (2015).
25. Qiu, X. *et al.* Pien Tze Huang alleviates relapsing-remitting experimental autoimmune encephalomyelitis mice by regulating Th1 and Th17 cells. *Front. Pharmacol.* **9**, 1237 (2018).
26. Botia, J. A. *et al.* An additional k-means clustering step improves the biological features of WGCNA gene co-expression networks. *BMC Syst. Biol.* **11**, 47 (2017).
27. Wan, Q., Tang, J., Han, Y. & Wang, D. Co-expression modules construction by WGCNA and identify potential prognostic markers of uveal melanoma. *Exp. Eye Res.* **166**, 13–20 (2018).
28. Han, X. *et al.* Lineage tracing reveals the bipotency of SOX9(+) hepatocytes during liver regeneration. *Stem Cell Rep.* **12**, 624–638 (2019).
29. Inagaki, Y. & Okazaki, I. Emerging insights into transforming growth factor beta Smad signal in hepatic fibrogenesis. *Gut* **56**, 284–292 (2007).
30. Wu, L. *et al.* Quercetin prevents hepatic fibrosis by inhibiting hepatic stellate cell activation and reducing autophagy via the TGF-beta/Smads and PI3K/Akt pathways. *Sci. Rep.* **7**, 9289 (2017).
31. Laurent, M. A., Bonnier, D., Theret, N., Tuffery, P. & Moroy, G. In silico characterization of the interaction between LSKL peptide, a LAP-TGF-beta derived peptide, and ADAMTS1. *Comput. Biol. Chem.* **61**, 155–161 (2016).
32. Bukong, T. N., Maurice, S. B., Chahal, B., Schaeffer, D. F. & Winwood, P. J. Versican: A novel modulator of hepatic fibrosis. *Lab. Invest.* **96**, 361–374 (2016).
33. Xu, H. *et al.* RAP-8 ameliorates liver fibrosis by modulating cell cycle and oxidative stress. *Life Sci.* **229**, 200–209 (2019).
34. Lin, I. Y. *et al.* CCM111 prevents hepatic fibrosis via cooperative inhibition of TGF-beta, Wnt and STAT3 signaling pathways. *J. Food Drug Anal.* **27**, 184–194 (2019).
35. Weerawatanakorn, M. *et al.* Protective effect of theaflavin-enriched black tea extracts against dimethylnitrosamine-induced liver fibrosis in rats. *Food Funct.* **6**, 1832–1840 (2015).
36. Moreira, S. T. *et al.* Influence of cytokine and cytokine receptor gene polymorphisms on the degree of liver damage in patients with chronic hepatitis C. *Meta Gene.* **9**, 90–96 (2016).
37. Rubbia-Brandt, L. *et al.* Gene expression profiling provides insights into pathways of oxaliplatin-related sinusoidal obstruction syndrome in humans. *Mol. Cancer Ther.* **10**, 687–696 (2011).
38. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods.* **9**, 671–675 (2012).
39. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014).
40. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
41. Kovaka, S. *et al.* Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* **20**, 278 (2019).
42. Tremblay, B. L., Guenard, F., Lamarche, B., Perusse, L. & Vohl, M. C. Weighted gene co-expression network analysis to explain the relationship between plasma total carotenoids and lipid profile. *Genes Nutr.* **14**, 16 (2019).
43. Zhu, J. *et al.* Identification of tissue-specific protein-coding and noncoding transcripts across 14 human tissues using RNA-seq. *Sci. Rep.* **6**, 28400 (2016).
44. Shannon, P. *et al.* Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **13**, 2498–2504 (2003).

## Acknowledgements

The project design, materials, RNA-seq data production, writing and publication fee were supported by Grants to Shengying Qin as follows: the 863 program (2012AA02A515, 2012AA021802), the National Nature Science Foundation of China (81773818, 81273596, 30900799, 81671326), the National Key Research and Development Program (2017YFC0909303, 2016YFC0905000, 2016YFC0905002, 2016YFC1200200), the 4th Three-year Action Plan for Public Health of Shanghai (15GWZK0101), the Shanghai Pujiang Program (17PJJD020) and the Shanghai Key Laboratory of Psychotic Disorders (13dz2260500). The RNA-seq data analysis and interpretation was supported by grants to Jinhang Zhu awarded by Anhui Medical University for Scientific Research of BSKY (XJ201607) and Anhui Medical University for Scientific Research (2017xkj006). The PZH drug acquisition was supported by grants from the Natural Science Foundation of Fujian Province (2016J05210) of Fei Hong. The funder provided support in the form of salaries for Zhiliang Chen and Fei Hong, but did not have any additional role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript. The specific roles of these authors are articulated in the ‘Author contributions’ section.

## Author contributions

Project design: S.Y.Q. and J.H.Z. Experiments and data analysis: J.H.Z., D.Z., T.W. and H.W. Preparation of materials: C.H., N.Z., M.Y.W. and J.S. Manuscript preparation: J.H.Z., D.Z. and L.C. PZH drug acquisition: F.H. and Z.L.C., All authors read and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-96459-5>.

**Correspondence** and requests for materials should be addressed to F.H. or S.Q.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021