Cancer Management and Research

Elevate Series

Open Access Full Text Article

ORIGINAL RESEARCH

Identification of Candidate Hub Genes and Drug Targets for Cholangiocarcinoma via Juhua (Chrysanthemum Morifolium) Bioactivity and Molecular Docking: A Bioinformatics Approach

Song Yang, Wan-Liang Sun, Shuo Zhou, Zheng Lu

The First Affiliated Hospital of Bengbu Medical College, Bengbu, 233004, People's Republic of China

Correspondence: Zheng Lu, The First Affiliated Hospital of Bengbu Medical College, Bengbu, 233004, People's Republic of China, Email luzhengdr@163.com

Background: Cholangiocarcinoma (CHOL) is a malignancy with poor clinical outcomes and limited treatment options. While extensive research has investigated genetic and signaling pathways in CHOL, the molecular mechanisms underlying disease pathogenesis remain incompletely understood. A key hurdle has been the lack of a systematic, multi-omic approach to illuminate causal relationships between genetic variants and CHOL risk.

Results: We integrated gene expression, co-expression network, and Mendelian randomization analyses to elucidate molecular drivers of CHOL. Gene set enrichment of differentially expressed genes from CHOL tumor samples identified enrichment in cancer-related biological processes. Weighted gene co-expression network analysis identified modules highly correlated with CHOL, including genes involved in cell cycle regulation, transcription, and proteolysis. Integrating these data with targets of the herbal formula Juhua, which shows anti-CHOL activity, pinpointed four candidate hub genes (CDK5, CDK7, CTSB, MAP2K2). Molecular docking revealed interactions between Juhua constituents and these hub genes. Mendelian randomization analysis of genetic variants implicated CCL2, CD5, CXCL6, CXCL9, HGF, IL10, IL10RA, IL18, IL24, IL2RB, IL6, IL8, SIRT2 and SLAMF1 as causally associated with CHOL.

Conclusion: Our multi-omic analysis provides new insight into molecular underpinnings of CHOL and identifies candidate disease drivers, signaling pathways and herbal targets for further validation. This systematic approach established a framework for illuminating causal links between genetics, molecular mechanisms and disease pathogenesis, with potential to accelerate drug and biomarker development for CHOL.

Keywords: bioinformatics, cholangiocarcinoma, Mendelian randomization, Juhua

Introduction

Cholangiocarcinoma (CHOL) is a rare but aggressive malignancy originating from the epithelial cells lining the bile ducts. Its low incidence, diverse clinical presentations, and lack of early diagnostic markers contribute to significant challenges in management and treatment. Despite advances in medical technologies, the prognosis for CHOL remains dismal, with a 5-year survival rate of less than 10%.[1,2] Current research efforts are focused on early detection and the development of more effective therapeutic strategies to improve patient outcomes.

The pathogenesis of CHOL is complex, involving the malignant transformation of bile duct epithelial cells due to genetic mutations and environmental factors. Chronic inflammation, liver dysfunction, and bile stasis have been implicated as key risk factors, promoting DNA damage and the proliferation of malignant cells.[3] CHOL often manifests with bile duct obstruction, leading to symptoms such as jaundice and liver dysfunction, but diagnosis typically occurs in advanced stages when the tumor has already metastasized.[4] Thus, early detection and intervention are critical for improving survival rates.

Traditional Chinese medicine (TCM), particularly herbs like Chrysanthemum morifolium (Juhua), has been widely used in liver-related conditions. Chrysanthemum, a genus known for its diverse phytochemical compounds, has been

Received: 13 August 2024
Accepted: 15 November 2024
Published: 9 December 2024

investigated for its potential therapeutic properties, including its relevance to CHOL. Research has identified several compounds within Chrysanthemum that exhibit anticancer activities, making it a promising candidate for CHOL treatment. For instance, luteolin, a flavonoid found in Chrysanthemum, has been shown to inhibit the Wnt/β-catenin signaling pathway by interacting with β-catenin, leading to cell cycle arrest and apoptosis in cholangiocarcinoma cells. Luteolin's ability to alter the structural conformation of β-catenin and reduce its nuclear accumulation underscores its potential as a therapeutic agent.[5] Quercetin, another major flavonoid, is known for its antioxidant properties and its ability to inhibit the proliferation and invasion of intrahepatic cholangiocarcinoma (ICC) cells.[6] It promotes apoptosis and G1-phase cell cycle arrest while inducing ferroptosis through the nuclear factor-kappa-B (NF-κB) pathway, suggesting a dual mechanism for combating cancer growth.[7] Similarly, kaempferol has demonstrated the ability to induce G2/M phase cell-cycle arrest in CCA cells and synergizes with Smac mimetics to induce necroptosis, specifically targeting tumor cells while sparing non-tumor cells.[8] These bioactive compounds provide a scientific basis for further exploring the therapeutic potential of Chrysanthemum in cholangiocarcinoma.[9]

To better understand how Juhua and its compounds influence cholangiocarcinoma, a multi-approach strategy that integrates various bioinformatics techniques—network pharmacology, weighted gene co-expression network analysis (WGCNA), and Mendelian randomization (MR)—is crucial.[10] Each of these methods offers specific strengths that, when combined, provide a more comprehensive understanding of the complex interactions between herbal compounds and disease pathways.

Network pharmacology allows for the identification of multiple molecular targets and pathways associated with cholangiocarcinoma. Instead of focusing on a single target, this approach reveals how Juhua's key compounds, such as luteolin, quercetin, and kaempferol, may influence different cellular processes simultaneously, offering insights into potential synergistic effects and broader therapeutic mechanisms.

WGCNA helps identify clusters of co-expressed genes and pinpoint hub genes that play central roles in cholangiocarcinoma progression. By analyzing gene networks modulated by Juhua's bioactive compounds, WGCNA provides a functional context for how these compounds influence disease-relevant pathways.

Mendelian randomization (MR) strengthens these findings by using genetic variants as instrumental variables to infer causal relationships. This allows us to determine whether the genetic effects linked to Juhua's compounds directly contribute to cholangiocarcinoma development, offering a robust layer of validation for the bioinformatics predictions.

The advantage of this multi-method strategy lies in its ability to capture both the breadth and depth of Juhua's potential effects on cholangiocarcinoma. While network pharmacology identifies broad molecular interactions, WGCNA refines this by linking these interactions to key gene networks. MR further validates the causal relevance of these findings. Relying on a single method would provide limited insights and could miss important interactions or pathways. By combining these approaches, this study not only uncovers new therapeutic targets but also enhances the precision and reliability of the results, offering more effective strategies for improving cholangiocarcinoma prognosis.

In summary, this multi-approach analysis uncovers novel targets and pathways influenced by Juhua and paves the way for developing more precise treatment strategies. Experimental validation of these predictions is essential to confirm their relevance, ultimately advancing the use of Juhua in CHOL therapy and drug development.

Materials and Methods

Data Sources and Processing

We utilized Gene Expression Omnibus (GEO) to obtain GSE26566 for our analysis. The datasets were derived from Illumina HumanRef-8 v2.0 expression beadchip and consisted of 104 freshly-frozen tumor samples and 59 matched non-cancerous liver samples.[11] To identify differentially expressed genes (DEGs), we employed the "Limma" and "edgeR" packages in R software, as previously described. DEGs were defined as those with P-adjusted values less than 0.05 and a log2 (fold change) greater than 1 or less than −1. By implementing this methodology, we were able to effectively identify genes that were differentially expressed between tumor and non-cancerous liver samples.[12]

GO and KEGG Pathway Enrichment Analysis

To gain further insights into the biological functions and pathways associated with DEGs identified in our study, we performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis.[13] We conducted functional enrichment analysis using the "ClusterProfiler" package in R and analyzed the results for GO and KEGG pathway enrichment. GO and KEGG terms with a P-value less than 0.01 were considered to be significantly enriched. The GO categories were divided into biological processes (BP), molecular functions (MF), and cellular components (CC), while the KEGG pathway analysis provided insights into the signaling pathways and networks. By performing this functional enrichment analysis, we were able to gain valuable insights into the biological functions and pathways that were associated with the DEGs of interest.

Weighted Gene Co-Expression Network Analysis (WGCNA)

Using the scale-free topology criterion, we constructed a co-expression network in the GSE26566 dataset with the aim of identifying co-expression gene modules. This was done using the "WGCNA" package in R software. WGCNA began by selecting the top 5000 variants. Pearson's correlations between gene pairs constructed an adjacency matrix. A soft threshold power (β) ensured a biologically meaningful scale-free network topology. Dynamic Tree Cut identified gene modules. Hub genes were designated based on the top three highest Pearson module membership correlation values, with $P<0.05$.[14]

Screening for Targets of Juhua and Cholangiocarcinoma (CHOL)

We aimed to identify the main components of Juhua and their potential protein targets, as well as key genes associated with CHOL regulation. To accomplish this, we utilized a combination of traditional Chinese medicine systems pharmacology (TCMSP) analysis and bioinformatic databases.[15] First, we screened the main components of Juhua using the TCMSP database, based on criteria such as drug-like properties (DL) and oral bioavailability (OB%). Compounds with OB% greater than or equal to 30% and DL greater than or equal to 0.18 were considered potential candidates for further analysis. Next, we utilized SwissTargetPrediction, SuperPred, and SEA to identify the potential protein targets of the 20 active compounds in Juhua.[16]

To identify key genes associated with CHOL regulation, we searched for CHOL-related targets in the GeneCards database, the OMIM database, and the DisGeNET database.[17] We then identified candidate key genes by taking the intersection of Juhua-related targets, DEGs, key module genes, and CHOL-related targets. To visualize the results, we utilized Venn diagrams.

Molecular Docking

To investigate the interaction activity between key genes and their corresponding active ingredients, we employed molecular docking using AutoDock Vina. The 2D structure of the active ingredient was obtained from TCMSP, with hydrogen atoms added using AutoDock Vina. Key target proteins were obtained from the Protein Data Bank (PDB), separated from their ligands, and hydrogen atoms were added. The charge was calculated, and the size and center of the docking box were determined. We then utilized Vina to dock the active ingredients with the target protein to evaluate the binding ability of the molecules. A binding energy of less than zero suggests that the ligand can bind to the receptor instinctively.[18] The lower the energy score of the ligand-receptor binding configuration, the higher the probability of binding.[19] Visualization was achieved through Vina and Logplot+.

In vitro Experiment

Cell Culture and Extraction of Juhua

The human cholangiocarcinoma cell lines, HuCC-T1, obtained from Cellcook (Guangzhou, China), were cultured and maintained in a humidified atmosphere containing 5% CO2 at 37°C in RPMI 1640 (GIBCO, Life Technologies, Grand Island), supplemented with 10% fetal bovine serum (GIBCO).

Air-dried Juhua was treated twice with 80% aqueous ethanol (1:20 w/v), mixed, and sonicated at 60°C for 40 minutes. The resulting solution was then concentrated using a rotary evaporator, and was purified using the D-101 macroporous resin. The concentration of Juhua in the final extract was determined to be 50.36% using a UV-vis spectrophotometer, with rutin serving as the reference compound.

Measurement of Cell Growth and Viability

A Cell Counting Kit-8 (CCK-8) was purchased from Dojindo Laboratories (Kumamoto, Japan) and used to evaluate cell proliferation, according to the manufacturer's instructions. Briefly, 1×10^5 cells were seeded into each well of a 96-well plate and cultured in 100 μL of RPMI-1640 supplemented with 10% FBS. After 24 h, Juhua (0, 0.01, 0.05, 0.1, or 0.5, 1 mg/mL) was added to each well, and the cells were cultured for an additional 72 h. CCK-8 reagent (10 μL) was added to each well, and the plates were incubated at 37°C for 3 h. The absorbance of each well was measured at 450 nm using a microplate reader.

Cell Treatment

The cells were continuously cultured for 4 days and then treated with normal culture medium or culture medium containing different concentrations of Juhua (10% fetal bovine serum). Based on the previous cell viability assay experiment, the selected concentrations are 0, 0.01, 0.05, 0.1, and 0.5 mg/mL. After 24 hours of treatment, the cell samples were collected.

Real-Time Reverse Transcription-PCR

Total RNA was isolated from cells using the Trizol Total RNA Isolation kit (Beyotime, Shanghai) according to the manufacturer's protocol. RNA was eluted with RNase-free water. Reverse transcription (RT)-PCR was performed using the Transcriptor First Strand cDNA Synthesis kit (Roche Molecular Biochemicals, Indianapolis) according to the manufacturer's protocol. Briefly, reactions were incubated at 65 °C for 10 min, at 55 °C for 30 min, and then at 85 °C for 5 min. Oligonucleotide primers (Table 1) were designed using the Primer 5 software and synthesized by Invitrogen. Real-time monitoring of PCR products was carried out using the SYBR Green master mix (Roche Molecular Biochemicals) and the Prism 7500 Real Time PCR Detection System (Applied Biosystems, Foster City). Cycling conditions were 95 °C for 10 min, followed by 40 repeats of 95 °C for 15s and 60 °C for 1 min. Levels of mRNAs were calculated using the comparative cycle threshold (ΔΔCT) method and normalized to the levels of GAPDH, the internal control, to obtain the relative mRNA level of each target.

Mendelian Randomization Analyze

To investigate the relationship between the relevant genes and CHOL, we initially utilized the target genes identified and subsequently employed the Mendelian Randomization (MR) method to assess the association between these target genes and the disease. MR is a statistical approach that facilitates the evaluation of causal relationships between exposure and outcomes. Our exposure data were sourced from the UK Biobank (https://www.ukbiobank.ac.uk), specifically from the finn-b-C3 sample dataset, which represents a comprehensive, large-scale national biorepository. In our Mendelian randomization analysis, we employed the TwoSampleMR (https://github.com/MRCIEU/TwoSampleMR) package. We set the threshold for the minimum LD R2 value at 0.8 to ensure the robustness of the instrumental variables. For aligning palindromes, a minor allele frequency (MAF) threshold of 0.3 was established. We conducted multiple MR analyses

Table 1 Primer Sequences for the Genes CDK5R1, CDK7, CTSB, MAP2K2

Gene Name	Forward Primer	Reverse Primer
CDK5R1	5'-TCATCTCCGTGCTGCCTTGGAA-3'	5'-CTCATTGTTGAGGTGCGTGATGT-3'
CDK7	5'-GCACACCAACTGAGGAACAGTG-3'	5'-AAGTCGTCTCCTGCTGCACTGA-3'
CTSB	5'-GCTTCGATGCACGGGAACAATG-3'	5'-CATTGGTGTGGATGCAGATCCG-3'
MAP2K2	5'-GTGGTCACCAAAGTCCAGCACA-3'	5'-CACGATGTACGGCGAGTTGCAT-3'
GAPDH	5'-GTCTCCTCTGACTCAACAGCG-3'	5'-ACCACCCTGTTGCTGTAGCCAA-3'

using various methods, including Wald ratio, MR Egger, Weighted median, Inverse variance weighted, and Weighted mode. To mitigate potential biases and enhance the validity of our results, we rigorously excluded SNPs in linkage disequilibrium.

Statistical Analysis

All statistical analyses in our study were implemented using R software (version 4.2.2). The difference between the two groups was analyzed by Student's *t*-test. All statistical P-values were two-sided, and statistical significance was considered with P-value < 0.05.

Results

Identification of DEGs and Functional Enrichment Analysis

A comprehensive analysis of the GSE26566 dataset revealed a total of 942 differentially expressed genes (DEGs), consisting of 747 upregulated and 195 downregulated genes (Supplementary Table 1). The heatmap depicting the expression patterns of all genes and DEGs is presented in Figure 1A and B. To elucidate the biological functions of these DEGs, we conducted a functional enrichment analysis, which included KEGG enrichment analysis (Figure 1C and D). This analysis indicated the involvement of DEGs in multiple cancer-related pathways. While the study detected 942 DEGs, the genes differentially expressed specifically in cholangiocarcinoma did not exhibit notable expression differences that were distinctly characteristic of this type of cancer, suggesting a complex genetic landscape that may not be fully represented by DEGs alone. Consequently, no significant enrichment clustering was observed within cancer-related categories, indicating the complexity of the genetic

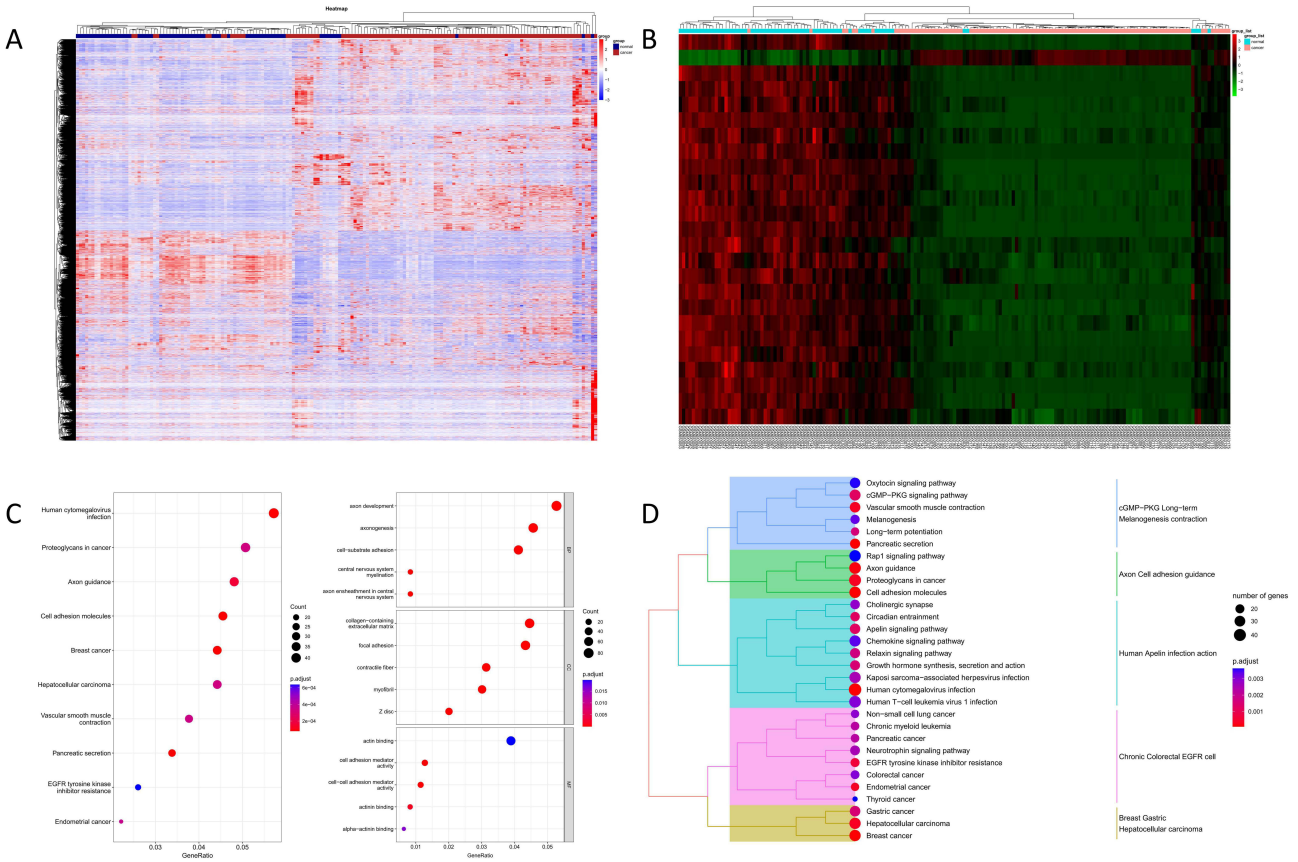

Figure 1 Functional analysis of differentially expressed genes (DEGs) in cholangiocarcinoma (CHOL). (**A**) Heatmap showing the expression of all genes in dataset GSE26566. (**B**) Heatmap displaying the expression patterns of the 2027 identified DEGs, including 1050 upregulated and 977 downregulated genes. (**C**) Bubble chart illustrating the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis for DEGs. (**D**) KEGG pathway classification of DEGs, highlighting pathways relevant to cancer and other biological functions.

changes in cholangiocarcinoma and the potential limitations of DEGs in capturing the full spectrum of signaling pathways and mechanisms specific to this disease. This lack of distinct enrichment clustering may reflect the inherent differences in genetic changes and molecular mechanisms between cholangiocarcinoma and other cancers, and suggests that DEGs alone may not comprehensively represent the complex signaling pathways and mechanisms specific to cholangiocarcinoma.

Construction of the Co-Expression Network by WGCNA in GSE26566

This study aimed to identify disease-associated genes and modules by analyzing gene expression data from patients with a specific condition. We started by selecting the top 5000 genes from the dataset based on their standard deviation. We then employed the "f1ashClust" algorithm to remove outliers from the dataset, retaining 36 samples for analysis. To ensure network scale-freeness, we applied the "pickSoftThreshold" function from the "WGCNA" package to screen soft thresholds ranging from 1 to 20 and selected a threshold of 6. We merged similar modules in the cluster dendrogram using a threshold of 0.20 (Figure 2A).Our analysis identified 16 modules with similar co-expressed genes (Figure 2B). The top 3 modules by absolute correlation were turquoise (r = −0.8, P=6e-39), green (r = 0.72, P=1e-28), brown (r = 0.53, P=2e-13) (Figure 2C). The turquoise (2287 genes), green (241 genes), and brown (335 genes) modules showed the strongest correlation with genes related to CHOL. The topological overlapping heat map depicted the TOM including all genes. An eigengene adjacency heatmap showed the correlation between different modules (Figure 2D). Top 25 heatmaps

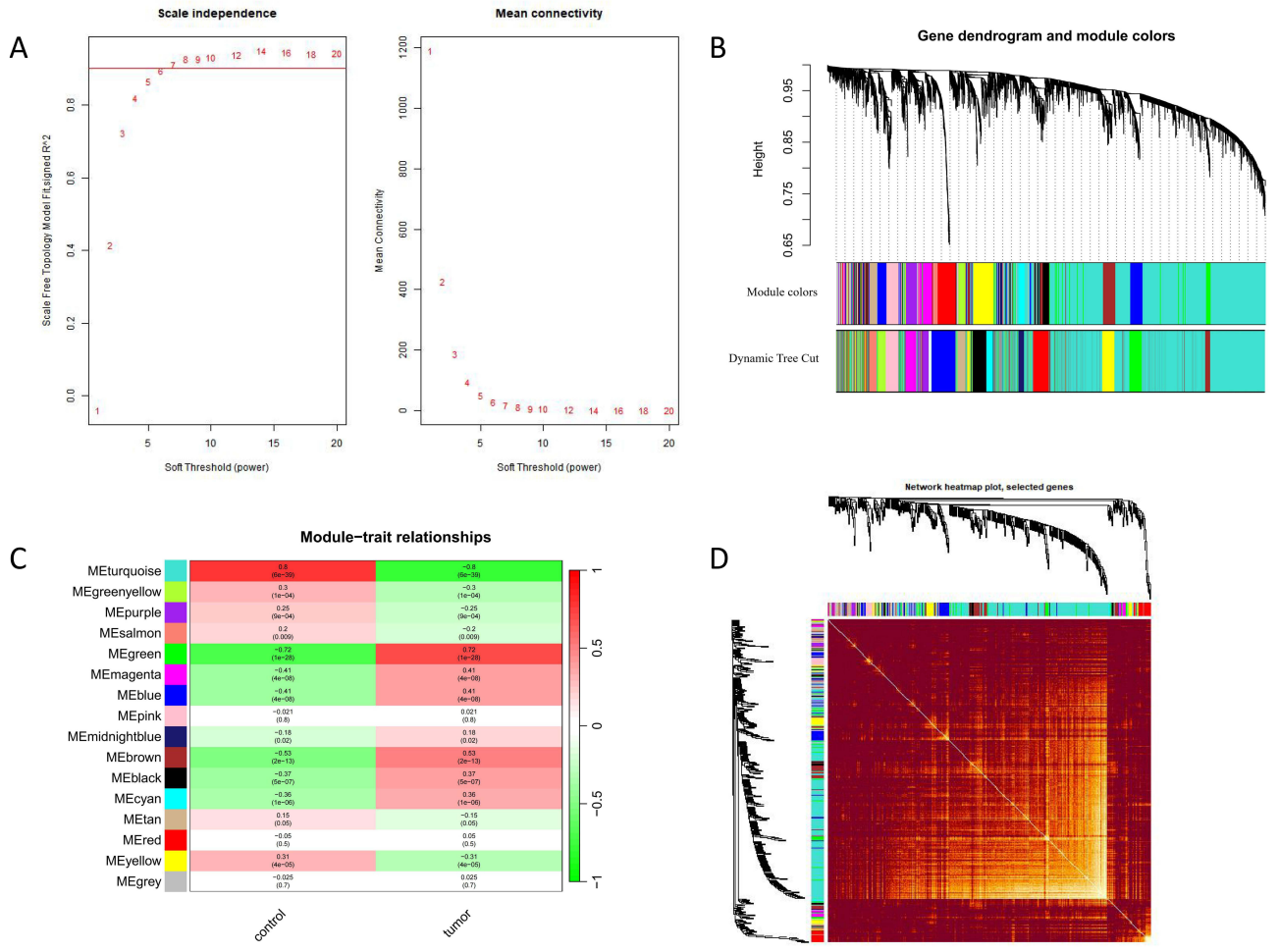

Figure 2 Weighted Gene Co-expression Network Analysis (WGCNA) for dataset GSE26566. (**A**) Scale-free fit index showing soft-threshold power selection to ensure network scale-freeness, with a chosen threshold of 6. (**B**) Mean connectivity plot confirming the selection of soft-threshold power to establish a robust network structure. (**C**) Heatmap of module-trait relationships, displaying the correlation between identified gene modules and CHOL-specific traits. Modules such as turquoise (r = −0.8), green (r = 0.72), and brown (r = 0.53) exhibit high correlations. (**D**) Clustering dendrogram of module genes with corresponding topological overlap matrix, identifying 16 modules with co-expressed genes potentially associated with CHOL.

of differentially expressed genes in these modules are shown in Figure 3. These findings suggest that the identified modules may play important roles in the development and progression of the studied disease. Further research could explore the specific functions of these genes and modules, as well as their interactions with other factors involved in the disease.

Identification of Candidate Hub Genes

We used a multi-faceted approach that involved integrating information from multiple sources. The 20 compounds Juhua are described in Supplementary Table 2. Specifically, we first utilized three target prediction database, namely SwissTargetPrediction, SuperPred, and SEA, to identify potential protein targets of the 20 active compounds in Juhua. After deduplication, we obtained a database of 325 targets that could potentially be modulated by the active compounds in Juhua.

We aimed to identify candidate hub genes that may be associated with CHOL. To achieve this goal, we turned our attention to CHOL-related targets and searched the GeneCards, OMIM, and DisGeNET databases for relevant information. After removing duplicates, we obtained a database of 2517 CHOL-related targets that could be involved in the regulation of CHOL levels.

Having obtained these databases, we intersected the differentially expressed genes in the GSE26566 dataset, the genes in the MEturquoise, MEgreen, and NEbrown modules, Juhua-related genes, and the CHOL-related genes to identify common genes that could potentially serve as candidate hub genes. After this analysis, we identified four genes as candidate hub genes for further investigation (Figure 4).

Overall, our multi-faceted approach allowed us to identify candidate hub genes that may be involved in the regulation of CHOL levels. Further studies are needed to confirm the role of these genes in CHOL regulation and to explore their potential as therapeutic targets for CHOL-related diseases.

Molecular Docking

In order to further investigate the potential binding activity of compounds with the identified hub genes, we performed molecular docking analysis. Specifically, we docked the top 20 compounds with CDK5, CDK7, CTSB, and MAP2K2. Our results showed that 97.5% of the binding energy data was less than −5 Kcal/mol, indicating strong binding activity between the receptor and ligand. The strongest binding energy observed was −10.04 Kcal/mol (Figure 5), suggesting a high degree of affinity between the ligand and receptor.

Among the top five docking results, CDK7-MOL001755, CDK5-MOL000358, CDK7-MOL000358, CDK7-MOL001790, and MAP2K2-MOL007326 demonstrated the highest binding affinities (Figure 6). We shifted our focus from describing amino acid residues to identifying the binding pockets, as these are more relevant to the functional

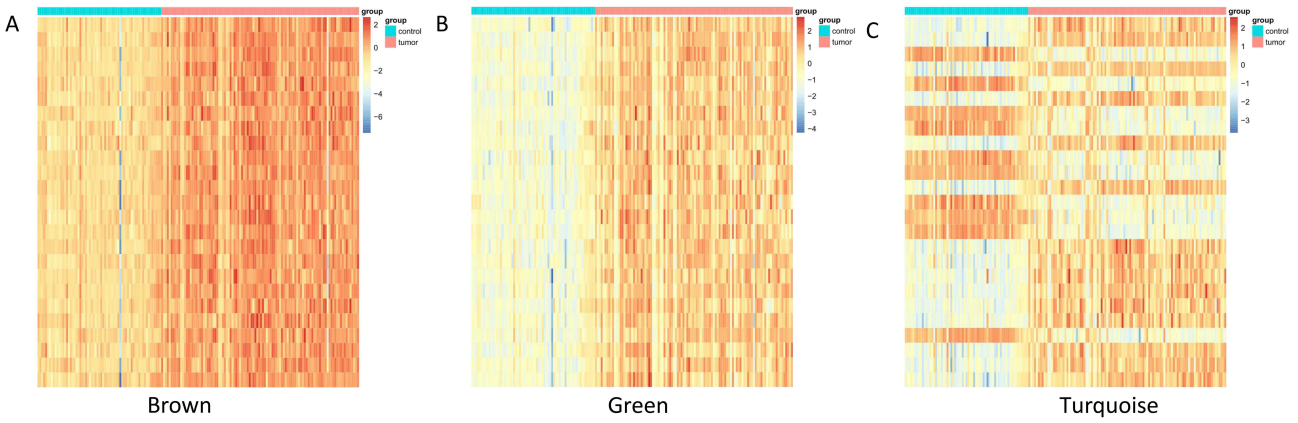

Figure 3 The top 25 heatmap of brown(**A**), green(**B**) and turquoise(**C**) genes. These modules were identified through the construction of a co-expression network using the WGCNA method applied to the GSE26566 dataset. The heatmaps visualize gene expression patterns within each module, highlighting how genes co-express with one another.

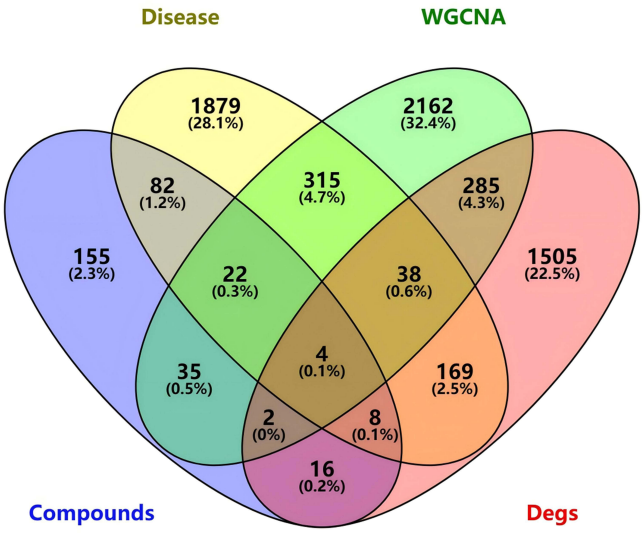

Figure 4 The Venny genes of Juhua anti CHOL.

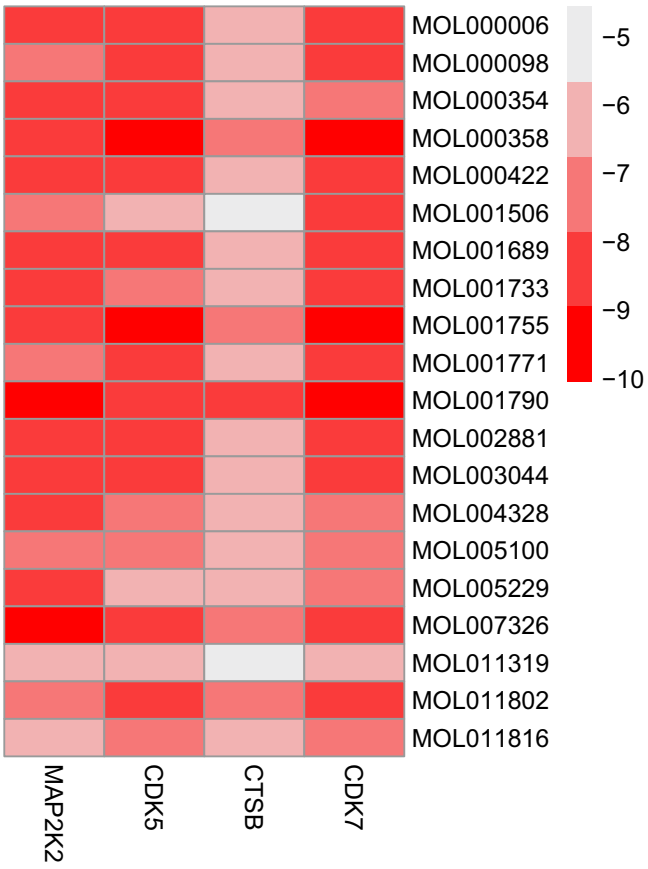

Figure 5 The heatmap of molecular docking.

activity of the proteins. The binding sites for these compounds were located on CDK7 and MAP2K2, and the specific interactions were identified through hydrogen bonding between the ligand and key residues within the binding pockets.

In particular, MOL007326 (Figure 6C) showed hydrogen bonding with ASP212, LYS196, GLYS4, ASN199, SER198, CYS211, and MET150 on MAP2K2, while MOL001790 formed hydrogen bonds with VAL314, GLU95,

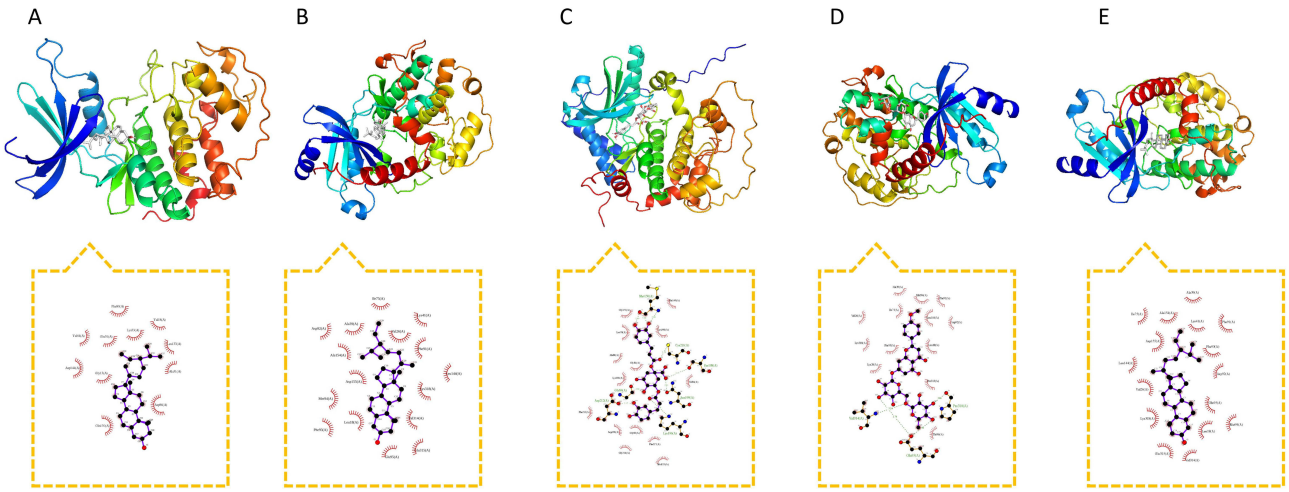

Figure 6 Molecular docking result. (**A**) CDK5-MOL000358; (**B**) CDK7-MOL000358; (**C**) MAP2K2-MOL007326; (**D**) CDK7-MOL001790; (**E**) CDK7-MOL001755.

and PRO310 on CDK7 (Figure 6D). These results suggest that these compounds may have potential therapeutic effects on CHOL-related diseases by modulating the activity of these hub genes through interactions within their binding pockets.

Overall, our molecular docking analysis provided insight into the potential binding activity of compounds with the identified hub genes and may serve as a starting point for further investigation into the development of novel therapeutic agents.

Juhua Affects Core Targets in HuCC-T1 Cells

In the assessment of HuCC-T1 cell growth and viability, we observed a significant decrease in cell viability of HuCC-T1 cells when the concentration of Juhua reached 1 mg/mL, as indicated by the CCK-8 assay (Figure 7A). Therefore, for subsequent RT-PCR experiments, we selected doses of 0.01, 0.05, 0.1, and 0.5 mg/mL, all expressed in the same unit of mg/mL for consistency.

Based on the above analysis, CDK5, CDK7, CTSB, MAP2K2 are cross targets. Therefore, we further examined the effect of Juhua on their mRNA expression in HuCC-T1 cells. As the dose of Juhua increased (Figure 7B), the mRNA expression of CDK5, CDK7, CTSB, MAP2K2 gradually decreased. At the dosages of 0.1 and 0.5 mg/mL, Juhua significantly reduced the mRNA expression of CDK5, CDK7, CTSB, MAP2K2.

Causal Effects of Target Genes on CHOL

SNPs: Basic Information

Cytokines served as the exposure factor, and CHOL was the outcome variable. A total of 50 cytokines were considered for exposure. SNPs were screened and identified as instrumental variables (IVs) with F values greater than 10 (Supplementary Table 3). The intercept of the MR-Egger regression, which is close to 0 ($P > 0.05$), suggests no horizontal pleiotropy of the IVs and minimal influence on the TSMR analysis results (Supplementary Table 4). Data with a significant Egger intercept, indicating potential pleiotropy or systematic bias, were excluded. Consequently, 23 cytokines were selected for TSMR analysis.

Two-Sample Mendelian Randomization Analysis

The MR results support a causal association between genetic susceptibility to specific cytokines and the risk of CHOL. In the absence of horizontal pleiotropy of IVs, inverse-variance weighted (IVW) was used as the primary method to estimate the causal association between genetic susceptibility to CHOL and the risk of cytokines. In the IVW, Weighted Median, and Simple Median modes, 15 cytokines were identified with a causal relationship with

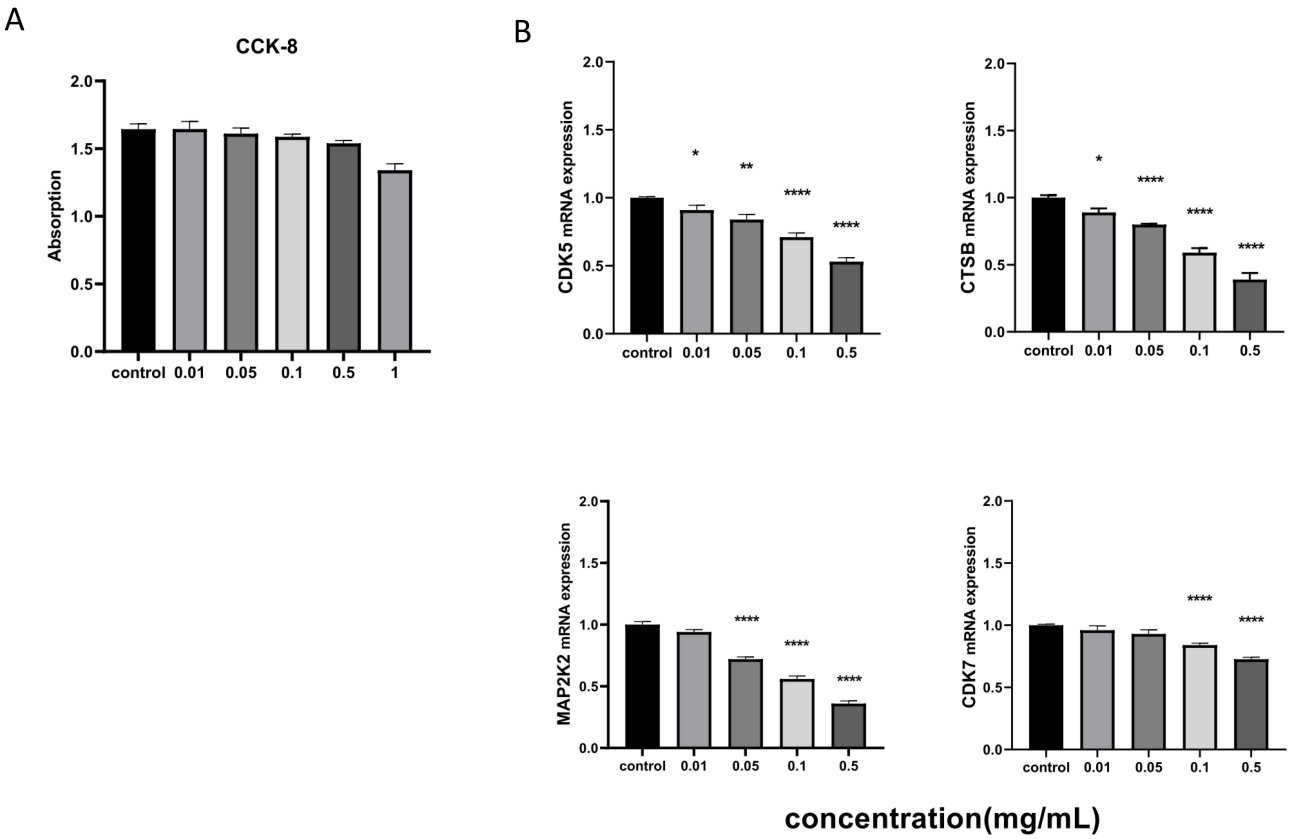

Figure 7 (**A**) The impact of different concentrations of Juhua on the viability of HuCC-T1 cells. (**B**)The mRNA expression of core targets on HuCC-T1 cells. *p < 0.05, **p < 0.01, ****p < 0.001, vs Control group.

CHOL. The specific data can be found in Supplementary Table 6. The detailed results are presented in Supplementary File 5 and summarized in Table 2 below.

Heterogeneity Test and Sensitivity Analysis

IVW and MR-Egger regression analyses were employed to detect heterogeneity between IVs. Heterogeneity was quantified using Cochran's Q test, with P < 0.05 indicating significant heterogeneity. Data with significant heterogeneity

Table 2 Cytokines with a Causal Relationship with CHOL

Cytokine	IVW OR (95% CI)	P-value
CCL2	0.10 (0.04–0.22)	1.490184e−08
CD5	0.87 (0.79–0.95)	2.830868e−03
CXCL6	0.15 (0.06–0.43)	3.738612e−04
CXCL9	3.10 (1.92–5.02)	3.978011e−06
HGF	0.30 (0.21–0.44)	5.251888e−10
IL10	1.90 (1.54–2.35)	5.251888e−10
IL10RA	0.77 (0.72–0.83)	5.384989e−14
IL24	0.36 (0.31–0.42)	1.826738e−39
IL2RB	0.42 (0.26–0.70)	8.808927e−04
IL6	2.07 (1.30–3.30)	2.255961e−03
IL8	1.94 (1.38–2.72)	1.353775e−04
LTA	0.24 (0.19–0.30)	6.334035e−33
MMP1	0.35 (0.27–0.46)	1.595187e−14
SIRT2	0.32 (0.14–0.72)	6.119563e−03
SLAMF1	0.87 (0.75–1.00)	4.715636e−02

were excluded. In Supplementary Table 7, significant heterogeneity was observed for the MMP1 gene, leading to its exclusion, and a total of 14 cytokines were retained in the final analysis.

Sensitivity analysis using the leave-one-out method was conducted to remove SNPs one by one, comparing the causal effects of the remaining SNPs with the TSMR analysis results of all SNPs to determine the robustness of the TSMR analysis results (Supplementary File 8). No sensitivity was observed in these 14 cytokines.

Discussion

In this study, our analysis of differentially expressed genes (DEGs) in the GSE26566 dataset identified 2027 genes with differential expression, but we did not observe clear enrichment clustering related to cholangiocarcinoma.[20] This may reflect the complex genetic mechanisms underlying cholangiocarcinoma and the limited understanding of this disease compared to other types of cancer.

To identify candidate hub genes that may be involved in cholangiocarcinoma (CHOL) regulation, we constructed a co-expression network using the WGCNA package and identified 16 modules with similar co-expressed genes. Among these modules, the turquoise, green, and brown modules showed the strongest correlation with CHOL-related genes.[21] These findings suggest that the identified modules may play important roles in the development and progression of CHOL-related diseases.

To further narrow down the candidate hub genes, we intersected the DEGs, module genes, Juhua-related genes, and CHOL-related genes and identified four genes as potential candidates: CDK5, CDK7, CTSB, and MAP2K2.[22] Molecular docking analysis revealed that these genes had strong binding activity with Juhua compounds, suggesting that they may be involved in the regulation of CHOL levels.

Binding pockets and protein activity: In our molecular docking analysis, we identified potential binding pockets for Linarin on CDK7 and Cynarin on MAP2K2. These binding pockets are crucial for protein activity as they can influence protein conformation and function.[23] By modulating these key residues, Linarin and Cynarin may alter protein conformation, thereby regulating the activity of CDK7 and MAP2K2 and affecting the expression of CHOL-related genes, including LDLR and ABCA1.[24] This regulation could significantly impact CHOL metabolism and homeostasis, offering potential therapeutic targets for CHOL-related diseases.

Integration of findings: To integrate the findings from different methods, we employed a comprehensive approach that included DEGs analysis, construction of co-expression networks, molecular docking analysis, and Mendelian randomization analysis. This multi-faceted strategy allowed us to identify key genes involved in CHOL metabolism and homeostasis and to elucidate their potential regulatory pathways.

Advantages of multi-approach strategy: The advantages of employing multiple approaches in our study are manifold. DEGs analysis helped us identify genes associated with cholangiocarcinoma, while co-expression network analysis revealed the interactions among these genes.[25] Molecular docking analysis provided potential binding affinities between these genes and Juhua compounds, and Mendelian randomization analysis offered evidence for causal relationships between genetic variants and CHOL. This multi-method strategy enhances the reliability and comprehensiveness of our findings.

Clinical relevance: Our findings contribute to cholangiocarcinoma treatment by identifying key genes that may be involved in the regulation of CHOL metabolism and homeostasis. The dysregulation of these genes has been implicated in various diseases, including cancer and neurodegenerative disorders. Further research into the roles of these genes in CHOL regulation and their potential as therapeutic targets may lead to new treatment strategies for patients with cholangiocarcinoma.

In summary, these four candidate hub genes, CDK5, CDK7, CTSB, and MAP2K2, have been shown to play important roles in various biological processes, and their dysregulation has been implicated in several diseases, including cancer and neurodegenerative disorders. The potential relationship between these genes and CHOL metabolism and homeostasis suggests that they may be involved in the regulation of CHOL-related pathways, including the sterol regulatory element-binding protein (SREBP) pathway, the low-density lipoprotein receptor (LDLR) pathway, and the ATP-binding cassette (ABC) transporter pathway. Further studies are needed to confirm the roles of these genes in CHOL regulation and to explore their potential as therapeutic targets for CHOL-related diseases.

Interestingly, our molecular docking analysis also revealed potential binding activity of Linarin with CDK7 and Cynarin with MAP2K2. Specifically, Linarin formed hydrogen bonding with key residues on CDK7, while Cynarin showed potential binding activity with MAP2K2.

The mechanism by which Linarin and Cynarin exert their effects on CDK7 and MAP2K2, respectively, is still not fully understood. However, previous studies have suggested that these compounds may modulate the activity of these hub genes by regulating the expression of key CHOL-related genes, including LDLR and ABCA1. Further studies are needed to confirm the binding activity of Linarin and Cynarin with CDK7 and MAP2K2, as well as to explore the underlying mechanisms by which these compounds modulate CHOL metabolism and homeostasis. Nonetheless, our findings suggest that Linarin and Cynarin may have potential therapeutic effects on CHOL-related diseases by targeting these hub genes.

In order to gain deeper insights into the biological mechanisms of CHOL, we conducted a Mendelian randomization analysis to analyze and assess target genes associated with CHOL. Our utilization of this approach aimed to elucidate the causal relationships between genetic variants and CHOL, providing a robust framework for understanding the molecular underpinnings of this condition.

Most genes associated with CHOL did not appear to have a significant statistical correlation with finn-b-C3 samples.[25] This might suggest that the role of these genes in the biological processes related to CHOL does not directly relate to the physiological or disease states represented by finn-b-C3 samples. However, several genes, such as CCL2, CD5, CXCL6, CXCL9, HGF, IL10, IL10RA, IL18, IL24, IL2RB, IL6, IL8, SIRT2, and SLAMF1, were identified to have a statistically significant association with the finn-b-C3 samples. These genes might serve as critical links between the biological mechanisms of CHOL and the physiological or disease states represented by the finn-b-C3 samples. It's worth noting that, although our analysis provides compelling evidence of the association between these genes and the finn-b-C3 samples, further experimental validation is required to ascertain whether these observed associations have biological relevance.

Conclusion

In conclusion, our study has made several significant contributions to the understanding of the genetic mechanisms underlying cholangiocarcinoma (CHOL) and its regulation. Firstly, through the analysis of differentially expressed genes (DEGs) in the GSE26566 dataset, we identified 2027 genes that exhibited differential expression, providing a comprehensive view of the genetic landscape associated with CHOL. This analysis, while not yielding clear enrichment clustering, underscores the complexity of genetic mechanisms in CHOL and highlights the need for further exploration.

Secondly, by constructing a co-expression network and identifying modules with strong correlations to CHOL-related genes, we were able to narrow down potential candidate hub genes. The turquoise, green, and brown modules emerged as significant, suggesting their crucial roles in the development and progression of CHOL-related diseases.

Thirdly, the intersection of DEGs, module genes, Juhua-related genes, and CHOL-related genes led us to identify four potential candidate hub genes: CDK5, CDK7, CTSB, and MAP2K2. Our molecular docking analysis revealed strong binding activity between these genes and Juhua compounds, indicating their potential involvement in the regulation of CHOL levels.

Furthermore, our study emphasized the importance of combining multiple approaches to uncover novel genetic mechanisms and potential therapeutic targets. The integration of DEGs analysis, co-expression network construction, molecular docking analysis, and Mendelian randomization analysis not only enhanced the reliability of our findings but also provided a multi-dimensional understanding of CHOL regulation.

Lastly, our findings suggest that these candidate hub genes may play roles in CHOL metabolism and homeostasis, potentially impacting pathways such as the sterol regulatory element-binding protein (SREBP) pathway, the low-density lipoprotein receptor (LDLR) pathway, and the ATP-binding cassette (ABC) transporter pathway. This insight opens avenues for future research to confirm the roles of these genes in CHOL regulation and to explore their potential as therapeutic targets for CHOL-related diseases.

We believe that our comprehensive approach and the identification of these candidate genes provide a solid foundation for future studies to build upon, with the ultimate goal of improving treatment strategies for patients with cholangiocarcinoma.

Ethics Statement

The research did not involve any human participants or data. The experiments were conducted solely using cell lines provided by a commercial biotech company, without any direct involvement of human subjects. Using GSE26566 public database does not require additional ethics committee approval.

Disclosure

The authors report no conflicts of interest in this work.

References

1. Chen HD, Huang C-S, Xu Q-C, et al. Therapeutic targeting of CDK7 suppresses tumor progression in intrahepatic cholangiocarcinoma. *Int J Biol Sci*. 2020;16(7):1207–1217. doi:10.7150/ijbs.39779
2. Cheng W, Yang Z, Wang S, et al. Recent development of CDK inhibitors: an overview of CDK/inhibitor co-crystal structures. *European J Med Chem*. 2019;164:615–639. doi:10.1016/j.ejmech.2019.01.003
3. Gfeller D, Grosdidier A, Wirth M, et al. SwissTargetPrediction: a web server for target prediction of bioactive small molecules. *Nucleic Acids Res*. 2014;42:W32–W38. doi:10.1093/nar/gku293
4. Gong KQ, Mikacenic C, Long ME, et al. MAP2K2 delays recovery in murine models of acute lung injury and associates with acute respiratory distress syndrome outcome. *Am J Resp Cell Mol Biol*. 2022;66(5):555–563. doi:10.1165/rcmb.2021-0252OC
5. Yang H, Zhao Y, Song W, et al. The inhibition of β-catenin activity by luteolin isolated from paulownia flowers leads to growth arrest and apoptosis in cholangiocarcinoma. *Int J Biol Macromol*. 2024;254(Pt 1):127627. doi:10.1016/j.ijbiomac.2023.127627
6. Granato M, Rizzello C, Gilardini Montani MS et al . Quercetin induces apoptosis and autophagy in primary effusion lymphoma cells by inhibiting PI3K/AKT/mTOR and STAT3 signaling pathways. *J Nutr Biochem*. 2017;41:124–136.
7. Lomphithak T, Jaikla P, Sae-Fung A et al. Natural Flavonoids Quercetin and Kaempferol Targeting G2/M Cell Cycle-Related Genes and Synergize with Smac Mimetic LCL-161 to Induce Necroptosis in Cholangiocarcinoma Cells. *Nutrients*. 2023;15(14): 3090.
8. Intuyod K, Priprem A, Pairojkul C, et al. Anthocyanin complex exerts anti-cholangiocarcinoma activities and improves the efficacy of drug treatment in a gemcitabine-resistant cell line. *Int J Oncol*. 2018;52(5):1715–1726.
9. Chen X, Sun B, Zeng J, et al. Molecular mechanism of Spatholobi Caulis treatment for cholangiocarcinoma based on network pharmacology, molecular docking, and molecular dynamics simulation. *Naunyn Schmiedebergs Arch Pharmacol*. 2024;397(8):5789–5806.
10. Xu L, Gao X, Xing J et al. Identification of a necroptosis-related gene signature as a novel prognostic biomarker of cholangiocarcinoma. *Front Immunol*. 2023;14:1118816. doi:10.3390/ijms24065992
11. Shen H, Bai X, Liu J, et al. Screening potential biomarkers of cholangiocarcinoma based on gene chip meta-analysis and small-sample experimental research. *Front Oncol*. 2022;12:100140.
12. Chen Y, Liu D, Liu P et al. Identification of biomarkers of intrahepatic cholangiocarcinoma via integrated analysis of mRNA and miRNA microarray data. *Mol Med Rep*. 2017;15(3):1051–1056.
13. Xue TC, Zhang BH, Ye SL et al. Differentially expressed gene profiles of intrahepatic cholangiocarcinoma, hepatocellular carcinoma, and combined hepatocellular-cholangiocarcinoma by integrated microarray analysis. *Tumour Biol*. 2021;36(8):5891–5899.
14. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.
15. Ru J, Li P, Wang J, et al. TCMSP: a database of systems pharmacology for drug discovery from herbal medicines. *J Cheminform*. 2014;6:13.
16. Daina A, Michielin O, Zoete V, et al. SwissTargetPrediction: updated data and new features for efficient prediction of protein targets of small molecules. *Nucleic Acids Res*. 2019;47(W1):W357–W364.
17. Piñero J, Ramírez-Anguita JM+, Saüch-Pitarch J, et al. The DisGeNET knowledge platform for disease genomics: 2019 update. *Nucleic Acids Res*. 2020;48(D1):D845–D855.
18. 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(2): 455–461.
19. Laskowski RA, Swindells MB. LigPlot+: multiple ligand-protein interaction diagrams for drug discovery. *J Chem Info Model*. 2011;51 (10):2778–2786. doi:
20. Andersen JB, Spee B, Blechacz BR, et al. Genomic and genetic characterization of cholangiocarcinoma identifies therapeutic targets for tyrosine kinase inhibitors. *Gastroenterology*. 2012;142(4):1021–1031.
21. Zhong YJ, Luo XM, Liu F, et al. Integrative analyses of bulk and single-cell transcriptomics reveals the infiltration and crosstalk of cancer-associated fibroblasts as a novel predictor for prognosis and microenvironment remodeling in intrahepatic cholangiocarcinoma. *J Transl Med*. 2024;22:1 422.
22. Da Z, Gao L, Su G, et al. Bioinformatics combined with quantitative proteomics analyses and identification of potential biomarkers in cholangiocarcinoma. *Cancer Cell Int*. 2020;20:130.
23. Forli S, Huey R, Pique ME, et al. Computational protein-ligand docking and virtual drug screening with the AutoDock suite. *Nat Protoc*. 2011;11 (5):905–919.
24. Amiri Souri E, Laddach R, Karagiannis SN, et al. Novel drug-target interactions via link prediction and network embedding. *BMC Bioinformatics*. 2022;23(1):121.
25. Wang x, Zhu Z . A Mendelian randomization analysis reveals the multifaceted role of the skin microbiota in liver cancer. *Front Microbiol*. 2024;15:142212.