Contents lists available at ScienceDirect

Digital Chinese Medicine

journal homepage: http://www.keaipublishing.com/dcmed

Molecular features of traditional Chinese medicine syndrome evolution in chronic liver diseases: a dynamic network biomarker analysis

Qingqing Chen[a], Hua Zhang[b], Dong Guo[a], Hong Cai[c*], Yiyu Lu[a*]

a. Institute of Interdisciplinary Integrative Medicine Research, Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China

b. Key Laboratory of Liver and Kidney Diseases (Ministry of Education), Institute of Liver Diseases, Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine, Shanghai 201203, China

c. First Division of Hepatology Department, Xiamen Hospital of Traditional Chinese Medicine, Xiamen, Fujian 361015, China

ARTICLE INFO

Article history
Received 08 January 2026
Accepted 21 April 2026
Available online 25 June 2026

Keywords
Traditional Chinese medicine syndromes evolution
Chronic liver disease
Chronic hepatitis B
Liver cirrhosis
Hepatocellular carcinoma
Dynamic network biomarkers
Random forest
PI3K-AKT signaling pathway

ABSTRACT

Objective To elucidate the biological basis of traditional Chinese medicine (TCM) syndromes from the perspective of "same syndrome, different diseases" in patients with chronic hepatitis B (CHB), liver cirrhosis (LC), and hepatocellular carcinoma (HCC), thereby providing a complementary approach for the diagnosis and treatment of chronic liver diseases (CLD).

Methods To investigate the dynamic characteristics of TCM syndromes in CLD, transcriptomic profiling of peripheral blood mononuclear cells (PBMCs) was performed from patients with CHB, LC, or HCC presenting with three TCM syndromes: liver gallbladder dampness heat syndrome (LGDHS), liver depression spleen deficiency syndrome (LDSDS), and liver kidney Yin deficiency syndrome (LKYDS). These participants were recruited at Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine between August 1, 2018 and December 31, 2021. Differentially expressed genes (DEGs) were identified using the random variance model (RVM) *F* test with false discovery rate (FDR) correction. Principal component analysis (PCA) and unsupervised hierarchical clustering were applied to visualize sample grouping. Dynamic network biomarkers (DNB) analysis was employed to detect critical transition stages during syndrome evolution, followed by Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses to characterize the functional roles and pathway involvement of the DNB members. Random forest (RF) analysis and the area under the receiver operating characteristic (ROC) curves (AUC) were used to rank the importance of candidate genes. External validation was performed using microarray data from an independent CLD cohort (GSE89377) and RNA-seq data from The Cancer Genome Atlas Liver Hepatocellular Carcinoma (TCGA-LIHC) dataset. Additionally, reverse transcription quantitative polymerase chain reaction (RT-qPCR) was performed on an independent cohort of LC patients to validate the expression levels of the candidate genes.

Results The study included a total of 132 participants. DNB analysis identified LDSDS stage as a critical tipping point during TCM syndrome evolution across CHB, LC, and HCC. The phosphoinositide 3-kinase/protein kinase B (PI3K-AKT) signaling pathway was consistently enriched in the DNB analysis across all three types of CLD, suggesting its potential involvement in the critical transition of TCM syndromes. Among the 24 core DNB members of the

*Corresponding author: Yiyu Lu, E-mail: yiyulu@shutcm.edu.cn. Hong Cai, E-mail: cguan2004@126.com.

Peer review under the responsibility of Hunan University of Chinese Medicine.

DOI: 10.1016/j.dcmed.2026.05.005

Citation: CHEN QQ, ZHANG H, GUO D, et al. Molecular features of traditional Chinese medicine syndrome evolution in chronic liver diseases: a dynamic network biomarker analysis. Digital Chinese Medicine, 2026, 9(2): 241-256.

PI3K-AKT pathway, four genes—integrin subunit beta 1 (*ITGB1*), collagen type IV alpha 1 chain (*COL4A1*), collagen type IV alpha 2 chain (*COL4A2*), and DNA damage inducible transcript 3 (*DDIT3*)—were identified by RF analysis (Gini score > 1) and ROC analysis. ROC analysis demonstrated high discriminative ability for distinguishing LGDHS from LKYDS in CHB patients, with AUC of 0.789 1 for *ITGB1*, 0.707 0 for *COL4A1*, 0.714 8 for *COL4A2*, and 0.894 5 for *DDIT3*. In the independent CLD cohort (GSE89377), all four genes showed significant stepwise upregulation from normal to CHB, LC, and HCC (all $P < 0.05$). In the TCGA-LIHC dataset, their expression progressively increased with tumor stage. RT-qPCR validation in an independent LC cohort (30 LGDHS vs. 30 LKYDS) confirmed that *ITGB1*, *COL4A2*, and *DDIT3* were significantly upregulated in LKYDS compared with LGDHS (P = 0.015 2, 0.018 6, and 0.024 7, respectively), whereas *COL4A1* showed a non-significant upward trend (P = 0.120 1).
Conclusion This study introduces a novel approach to understanding the molecular features underlying TCM syndrome evolution in CLD. The PI3K-AKT pathway and four identified genes (*ITGB1*, *COL4A1*, *COL4A2*, and *DDIT3*) play crucial roles in the transition from excess (LGDHS) to deficiency (LKYDS) via the critical LDSDS stage. These findings offer potential quantitative biomarkers and therapeutic targets for TCM syndrome differentiation and may help arrest syndrome progression in CLD.

1 Introduction

Chronic liver diseases (CLD) are highly prevalent worldwide [1, 2]. Encompassing chronic hepatitis B (CHB), liver cirrhosis (LC), and hepatocellular carcinoma (HCC), CLD imposes a substantial global health burden [3]. Traditional Chinese medicine (TCM) offers distinct advantages in the treatment of CLD [4, 5]. TCM introduces two fundamental therapeutic concepts: preventive treatment before disease onset, and personalized intervention based on individual constitution, syndrome, and timing—principles that align closely with precision medicine [6, 7].

TCM syndromes reflect dynamic imbalances in the interaction between the human body and its environment [8, 9]. In CLD, syndrome evolution follows a predictable trajectory: from excess to excess-deficiency complex, and ultimately to deficiency. This progression mirrors the progressive decline of visceral function and the accumulation of pathological damage. Specifically, liver gallbladder dampness heat syndrome (LGDHS) represents the early excess state driven by pathogenic dampness and heat, whereas liver kidney Yin deficiency syndrome (LKYDS) characterizes the advanced deficiency state marked by depletion of Yin fluid and impairment of liver and kidney functions. Liver depression spleen deficiency syndrome (LDSDS) is defined by liver Qi stagnation and spleen insufficiency. Within TCM pathogenesis theory, LDSDS serves as a transitional hub between excess and deficiency syndromes, suggesting its pivotal role in syndrome evolution.

From a systems biology perspective, disease progression is not linear but involves critical transitions governed by dynamic molecular networks rather than individual molecules. This phenomenon of network-level abrupt change is highly consistent with the TCM concept of syndrome transformation. Nevertheless, the molecular mechanisms linking TCM syndrome transition theory to its underlying network dynamics remain largely unexplored.

Several studies have attempted to characterize TCM syndromes using molecular approaches. Machine learning models have been used to differentiate cold and hot syndromes [10], and network pharmacology has revealed genetic signatures associated with specific syndromes [11]. A study of 594 patients with CHB similarly demonstrated that excess-deficiency complex syndrome emerges at a more advanced stage than excess syndrome alone [12]. In our previous work, we found that distinct TCM syndromes have unique biological bases that can be distinguished by gene and protein expression profiles [13, 14]. However, these studies relied primarily on conventional differential expression analysis, which provides static snapshots and fails to capture the dynamic transitions inherent to TCM syndrome evolution. Consequently, the biological basis of “same syndrome, different diseases” remains poorly understood at the network level.

To address this gap, the dynamic network biomarker (DNB) approach was developed to detect early warning signals preceding critical transitions in complex biological system [15, 16]. Unlike static biomarkers, DNB identifies a group of molecules that exhibit sudden fluctuations and strong correlations at the tipping point, thereby enabling the prediction of imminent phase shifts. This approach is therefore particularly well-suited for investigating the dynamic evolution of TCM syndromes.

In this study, the DNB approach was applied to examine the transcriptome profiles of peripheral blood mononuclear cells (PBMCs) from patients with CHB, LC, or HCC, as well as from normal controls, all of whom presented with defined TCM syndromes. We aimed to characterize the dynamic molecular signatures underlying TCM syndrome progression from CHB to LC to HCC

following hepatitis B virus (HBV) infection, within the framework of "disease-TCM syndrome-prescription". Our findings may offer new insights into the molecular features and mechanisms driving TCM syndrome progression from both dynamic and network perspectives.

2 Materials and methods

2.1 Participants

Individuals were recruited at Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine between August 1, 2018 and December 31, 2021, from whom blood samples were collected. The study protocol was approved by the Institutional Ethics Committee of Shanghai University of Traditional Chinese Medicine (2017-557-40-01), and written informed consent was obtained from all participants. The trial was retrospectively registered at ClinicalTrials.gov (NCT03189992).

2.2 Diagnostic criteria

2.2.1 Western diagnostic criteria (i) Diagnostic criteria for CHB. The diagnosis of CHB required persistent positivity of hepatitis B surface antigen (HBsAg) for at least 6 months, detectable serum HBV DNA, and elevated alanine aminotransferase (ALT) levels exceeding the upper limit of normal (ULN) on at least two occasions within a 6- to 12-month period. The ULN for ALT was defined as 40 U/L [17].

(ii) Diagnostic criteria for LC. The diagnosis of LC was based on a combination of clinical, laboratory, imaging, and/or histological findings. Clinical evidence included signs of portal hypertension (e.g., ascites, splenomegaly, and esophageal or gastric varices), jaundice, or hepatic encephalopathy. Laboratory abnormalities comprised thrombocytopenia (platelet count < 100×10^9/L), elevated bilirubin, reduced serum albumin (< 35 g/L), and prolonged prothrombin time [international normalized ratio (INR) > 1.3]. Imaging findings consistent with cirrhosis on ultrasound, computed tomography (CT), or magnetic resonance imaging (MRI) included liver surface nodularity, blunt liver edges, caudate lobe hypertrophy, and/or signs of portal hypertension (splenomegaly and portosystemic collaterals). For non-invasive diagnosis, a liver stiffness measurement (LSM) >17.0 kPa on transient elastography (FibroScan) was considered indicative of cirrhosis in patients with chronic HBV infections, in accordance with Chinese guidelines [17].

(iii) Diagnostic criteria of HCC. In high-risk patients (e.g., those with HBV/HCV infection or cirrhosis), the diagnostic imaging criterion was the presence of non-rim arterial phase hyperenhancement followed by washout on portal venous or delayed phase on contrast-enhanced CT, MRI, or contrast-enhanced ultrasound. For lesions > 2 cm, one positive imaging technique was sufficient; for lesions ⩽ 2 cm, at least two positive techniques (or one technique plus a second confirmatory modality) were required. Serum alpha-fetoprotein (AFP) ⩾ 400 μg/L (after excluding other causes) was also considered diagnostic. Atypical or non-high-risk cases required biopsy confirmation [18].

2.2.2 TCM diagnostic criteria TCM syndrome classification was performed according to the Standards of Traditional Chinese Medicine Syndrome Differentiation for Viral Hepatitis [19].

(i) Diagnostic criteria for LGDHS. The diagnosis of LGDHS was based on a combination of primary and secondary symptoms. Primary symptoms included dark reddish urine, bitter taste in the mouth, and yellow, greasy tongue coating. Secondary symptoms comprised constipation, dry mouth with thirst but reluctance to drink, fatigue, and a rapid pulse. A diagnosis was confirmed if at least one primary symptom and two secondary symptoms were present, or if a typical tongue coating was observed together with any primary or secondary symptom.

(ii) Diagnostic criteria for LDSDS. LDSDS was diagnosed based on the presence of abdominal distension, poor appetite, fatigue, and yellowish skin. A yellow, greasy tongue coating was considered a key diagnostic feature. Patients typically presented with a sluggish or weak pulse. Additional symptoms, such as loose stools or diarrhea, indicated spleen involvement.

(iii) Diagnostic criteria for LKYDS. LKYDS was characterized by dizziness, tinnitus, dry mouth, and a red tongue with scanty coating. The pulse was often thin and rapid. Diagnosis was confirmed when the patient exhibited these typical symptoms together with specific findings, including a red tongue with scant coating and a rapid pulse. This syndrome reflects an imbalance between liver and kidney Yin, contributing to the overall disease pattern.

2.3 Inclusion and exclusion criteria

Inclusion criteria for research group: (i) age between 18 and 65 years; (ii) diagnosis of CHB, LC, or HCC according to the Western diagnostic criteria described in section 2.2; (iii) for patients with LC or HCC, HBV related etiology confirmed by positivity for HBsAg and detectable serum HBV DNA at enrollment; (iv) written informed consent was provided by all participants.

Exclusion criteria for research group: (i) age under 18 or over 65 years; (ii) non HBV viral hepatitis, fatty liver disease, or autoimmune liver diseases; (iii) serious comorbidities (e.g., heart, kidney, lung, endocrine, or metabolic diseases); (iv) pregnant or lactating; (v) mental health disorders.

Normal controls were individuals with no history of viral diseases, liver diseases, or other major chronic

illnesses, who were age- and gender-matched to the patient group.

2.4 Clinical information collection

Clinical information was collected during the initial visit by senior TCM specialists using a standardized diagnostic scale [20-22]. The specialists were blinded to patients' laboratory results. Collected data included demographic information (age, sex, and medical history), symptom questionnaires, and liver function tests. Each participant's syndrome classification was determined using the four TCM diagnostic methods: inspection, auscultation and olfaction, inquiry, and palpation. Tongue and pulse characteristics were evaluated, and a structured survey form was completed for expert review. The final syndrome classification was reached by consensus among the three specialists following independent assessments. All clinical data were retrieved from the hospital's electronic medical record system.

2.5 Sample collection and processing

2.5.1 Blood sample collection Whole blood was collected into sterile ethylenediaminetetraacetic acid (EDTA)-coated vacuum tubes. To minimize RNA degradation, all samples were processed within 2 h after collection. Serum was separated by sedimentation at 4 °C for 2 h, followed by centrifugation at 3 000 ×g for 15 min. The resulting supernatant was then transferred to 1.5 mL cryovials and stored at – 80 °C until RNA extraction. All cryovials were clearly labeled with a unique participant ID and date of collection.

2.5.2 RNA isolation PBMCs were isolated from blood samples using Ficoll-based density gradient centrifugation and subsequently stored at – 80 °C [23]. Total RNA was extracted from PBMCs using TRIzol® Reagent (Invitrogen, USA). Briefly, cells were lysed in TRIzol, mixed with chloroform, and centrifuged to separate phases. RNA was precipitated from the aqueous phase with isopropanol, washed with 75% ethanol, air-dried, and dissolved in RNase-free water. RNA concentration and purity were assessed using a NanoDrop ND-1000 spectrophotometer (NanoDrop, ND-1000). Only samples with an A260/A280 ratio between 1.8 and 2.1 were used for subsequent analyses.

2.5.3 Microarray detection Double-stranded cDNA was synthesized and labeled with Cy3 using the NimbleGen One-Color DNA Labeling Kit (Roche NimbleGen, USA) according to the manufacturer's protocol. Microarray hybridization was performed using the Roche 12×135K Array (Cat. No. A6484-00-01) following the manufacturer's instructions. After washing, the arrays were scanned using a microarray scanner (Axon GenePix 4000B, Molecular Devices, USA) at 5 μm resolution. Raw data were normalized using the robust multi-array average (RMA) method. Probe-level signals were summarized into gene expression values using NimbleScan software (version 2.5; Roche NimbleGen, USA), which performed background correction, quantile normalization, and probe-level summarization. Genes with expression signals below the threshold (minimum error call probes < 50.0) were excluded from further analysis.

2.6 DNB analysis

DNB analysis was performed to identify critical transition states, as previously described [24]. According to the DNB theory, when a biological system approaches a tipping point, a cluster of molecules (the DNB) exhibits three distinct properties: (i) a significant increase in the coefficient of variation (CV), indicating elevated fluctuation; (ii) a marked increase in the Pearson correlation coefficients (PCC) among cluster members; and (iii) a decrease in the PCC between the DNB members and other molecules outside the cluster. The DNB score (composite index, CI) was calculated as:

$$CI = CV_I \times \frac{PCC_I}{PCCo}$$

where CV_I is the average standard deviation of the DNB members, PCC_I is the average absolute PCC among DNB members, and $PCCo$ is the average absolute PCC between DNB members and other molecules.

DNB analysis was performed separately in four sample sets derived from the 36 transcriptomic profiles: (i) the multi-disease set, including all participants; (ii) CHB patients across the three TCM syndromes; (iii) LC patients across the three TCM syndromes; and (iv) HCC patients across the three TCM syndromes. For each set, DNB members were identified according to the three criteria described above. The resulting DNB groups were designated as DNB_{multi}, DNB_{CHB}, DNB_{LC}, and DNB_{HCC}, respectively.

2.7 Construction of the DNB-differentially expressed gene (DEG) association network

A genome-wide human gene interaction network was downloaded from the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database (version 11.5) as the background network. Only interactions with a combined confidence score ⩾ 0.7 were retained. All DNB members were then mapped onto this background network. Network visualization was performed using Cytoscape (version 3.9.1).

2.8 Random forest (RF) analysis

RF was performed to rank the contribution of each gene to discriminate between patients with three different TCM syndromes, thereby identifying potential biological

markers for syndrome differentiation. The RF analysis was conducted using the "randomForest" package [25] in R software (version 4.4.0; R Foundation for Statistical Computing). The number of decision trees (ntree) was set to 500, and the number of candidate variables randomly sampled at each split (mtry) was set to the square root of the total number of genes. Each tree was built using a bootstrap sample, and the remaining out-of-bag samples were used for internal validation. Gene importance was evaluated using the mean decrease in Gini impurity index. The top-ranked genes were selected for subsequent analysis.

2.9 Reverse transcription quantitative polymerase chain reaction (RT-qPCR)

Total RNA was extracted from PBMCs as described in Section 2.5. RNA was reverse-transcribed into complementary DNA (cDNA) using a reverse transcriptase (TOYOBO, Japan) with oligo(dT) and random primers in a 20 μL reaction volume, following the manufacturer's instructions.

RT-qPCR was performed using SYBR® Green Real-Time PCR Master Mix (TOYOBO, Japan) on a LightCycler® 480 System (Roche, Switzerland). Each 25 μL reaction contained 2 × SYBR Green Master Mix, 0.4 μmol/L of each forward and reverse primer, and 0.5 μL of template cDNA. The primer sequences for the target genes and the internal control β-actin (*ACTB*) are listed in Table 1. The cycling conditions were as follows: initial denaturation at 95 °C for 5 min; followed by 40 cycles of denaturation at 95 °C for 30 s, annealing at 54 °C for 30 s, and extension at 72 °C for 15 s; with a final fluorescence acquisition step at 83 °C for 1 s. A dissociation (melting) curve analysis was performed after each run to confirm amplification specificity. Each sample was run in duplicate. Relative gene expression was calculated using the comparative Ct ($2^{-\Delta\Delta Ct}$) method, normalized to *ACTB*. Negative controls (no template) were included in each run.

2.10 External datasets

To validate the expression patterns of core genes across disease stages, two independent public datasets were used. The GSE89377 dataset was downloaded from the Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/). It contains genome-wide expression profiles of 107 peripheral blood samples, including 13 normal controls, 20 CHB, 34 LC, and 40 HCC samples. The Cancer Genome Atlas Liver Hepatocellular Carcinoma (TCGA-LIHC) dataset was obtained from the Genomic Data Commons (GDC) data portal (https://portal.gdc.cancer.gov/), providing RNA-sequencing data and clinical stage information (stage I – IV) for 369 HCC patients.

2.11 Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis

GO biological process enrichment analysis and KEGG pathway enrichment analysis were performed to characterize the functional roles and pathway involvement of the identified DNB members, DEGs, and the DNB-DEG association network. The analyses were carried out using the clusterProfiler package (version 4.0) in R software (version 4.4.0). A hypergeometric test was applied, and adjusted P values < 0.05 were considered statistically significant. The enrichment results were visualized using the ggplot2 package.

Figure 1 illustrates the complete workflow of this study, from participant recruitment to data analysis.

Table 1 Primer sequences for target genes validated by RT-qPCR

Gene symbol	Forward primer (5′ → 3′)	Reverse primer (5′ → 3′)	Length (bp)
Integrin subunit beta 1 (*ITGB1*)	AGAGTGCCGTAACAACTG	CTGACAAGAGTAACCATCCT	269
Collagen type IV alpha 1 chain (*COL4A1*)	TACACCAGGGCTGATAGG	CCATCTCTTCCAGGAGAAC	147
Collagen type IV alpha 1 chain (*COL4A2*)	GATCCAAAGGTGCAGTGG	TTTAAGCCCAGGCTGTCC	167
DNA damage-inducible transcript 3 (*DDIT3*)	CTGCTTCTCTGGCTTGGCTGAC	GAGTCGCCTCTACTTCCCTGGTC	297
ACTB	ACAGAGCCTCGCCTTTGCCG	ACATGCCGGAGCCGTTGTCG	104

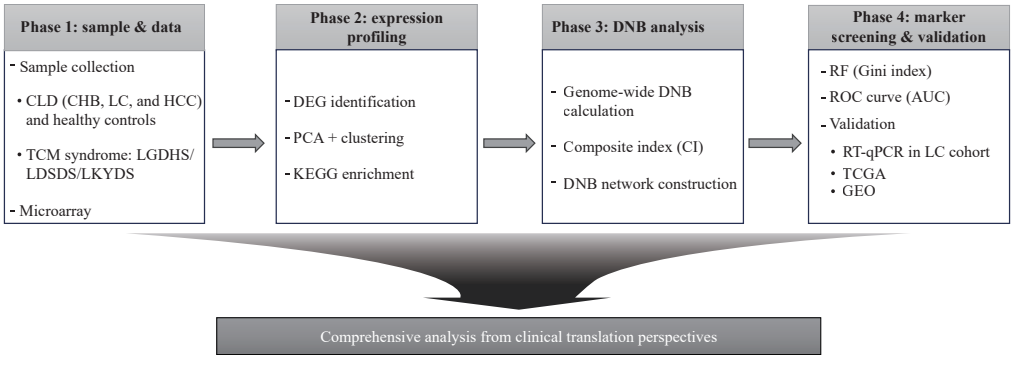

Figure 1 Schematic workflow of the study

2.12 Statistical analysis

DEGs were identified using the random variance model (RVM) *F* test, which improves statistical reliability for small sample sizes. Multiple testing correction was performed using the false discovery rate (FDR), and genes with an FDR < 0.05 were considered statistically significant. One-way analysis of variance (ANOVA) was also used to compare gene expression levels among multiple groups (healthy, LGDHS, LDSDS, and LKYDS) where applicable. Post-hoc comparisons were performed when appropriate. Principal component analysis (PCA) was performed to visualize the overall distribution of samples based on the identified DEGs. Unsupervised hierarchical clustering was conducted using the Euclidean distance metric and the complete linkage method to further assess sample grouping patterns. Both PCA and clustering analyses were carried out using R software (version 4.4.0) with the built-in functions prcomp and hclust, respectively. Receiver operating characteristic (ROC) analysis was performed to evaluate the discriminatory ability of each candidate gene. According to established criteria [26], an area under the ROC curve (AUC) $\geqslant$ 0.70 was considered indicative of acceptable diagnostic performance. The Mann-Whitney *U* test was used to compare gene expression levels between groups, with statistical significance set at $P < 0.05$. All statistical analyses were performed using R software (version 4.4.0; R Foundation for Statistical Computing, http://www.R-project.org).

3 Results

3.1 Transcriptomic profiling of TCM syndromes across CHB, LC, and HCC

A total of 132 eligible participants were recruited from Shuguang Hospital Affiliated to Shanghai University of Traditional Chinese Medicine. The mean age of the cohort was 53.5 ± 11.2 years, and 66.3% were male. Of the 132 participants, 36 underwent transcriptomic profiling by microarray, and the remaining 96 were reserved for independent RT-qPCR validation.

Genome-wide expression profiles of PBMCs were analyzed for these 36 participants, including 11 with LGDHS (4 CHB, 4 LC, and 3 HCC), 10 with LDSDS (4 CHB, 3 LC, and 3 HCC), 9 with LKYDS (including 3 CHB, 3 LC, and 3 HCC), and 6 normal controls. A comprehensive physical examination of TCM symptoms was conducted, and expression profiles of PBMCs were analyzed for participants with LGDHS, LDSDS, and LKYDS (Figure 2A). For example, among patients with CHB, the three TCM syndromes showed marked differences in tongue coating characteristics (Figure 2B). A total of 1 059 DEGs were identified using ANOVA among the four groups (healthy, LGDHS, LDSDS, and LKYDS) to characterize TCM syndrome-associated dysfunctions. PCA and unsupervised hierarchical clustering were then performed using these 1 059 DEGs (Figure 2C). Both analyses revealed three distinct clusters: (i) the healthy group, containing all control samples without the studied TCM syndromes and with a balanced constitution; (ii) the LGDHS group, encompassing most LGDHS samples, and (iii) the LKYDS group, including most LKYDS samples (Figure 2D). Notably, LDSDS samples did not form a separate cluster but instead intermingled with both LGDHS and LKYDS groups. Furthermore, samples with the same TCM syndrome were distributed across different disease stages, indicating that disease-associated differences did not drive the clustering results.

For Figure 2E, each TCM syndrome group was individually compared with the normal controls to obtain DEGs, which were then used for KEGG pathway enrichment analysis to characterize the unique functional disturbances associated with each TCM syndrome. Functional analyses revealed substantial overlap between the signaling pathways enriched by DEGs characterizing TCM syndromes in the multi-disease group and those characterizing the three disease stages. Dysfunction associated with LGDHS was predominantly enriched in cellular processes (e.g., the adipocytokine signaling pathway, protein digestion and absorption, and endocrine and other factor-regulated calcium reabsorption) and immune response-related signaling pathways [e.g., inflammatory mediator regulation of transient receptor potential (TRP) channels]. Dysfunctions associated with LDSDS were primarily enriched in metabolic processes (glutathione metabolism; glycine, serine, and threonine metabolism). The extracellular matrix (ECM)-receptor pathway was uniquely enriched in LDSDS. Dysfunctions associated with LKYDS were related to the biosynthesis of unsaturated fatty acids and protein biosynthesis [i.e., mechanistic target of rapamycin (mTOR) signaling pathway]. Notably, dysfunction of the phosphoinositide 3-kinase/protein kinase B (PI3K-AKT) signaling pathway was evident across both TCM syndrome progression and HCC development.

Conventional differential gene expression analysis provides a static overview of transcriptomic differences among groups, whereas DNB analysis further captures dynamic network changes to identify critical transition points during disease progression.

3.2 Identification of LDSDS as the critical tipping point in TCM syndrome evolution by DNB analysis

DNB analysis identified a strong signal of the crucial transition at the LDSDS stage (Figure 3A). To mitigate the influence of disease severity, the DNB method was applied to detect the critical stage in three specific diseases. Notably, the critical stage was consistently observed at LDSDS across CHB, LC, and HCC (Figure 3B). Subsequently, a series of molecular networks was constructed to visualize the correlations between DNB members and other

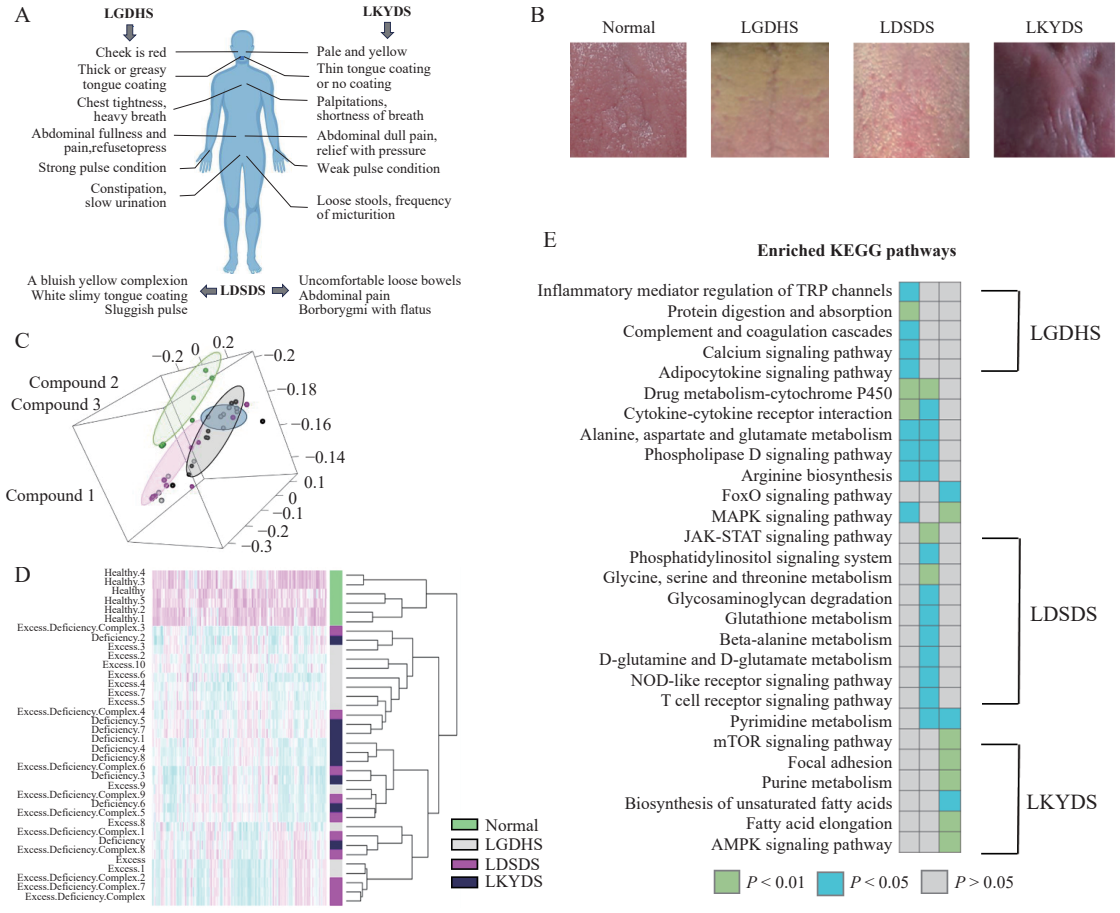

Figure 2 Integrated analysis of clinical symptoms and gene expression profiles of TCM syndromes

A, clinical symptoms used for TCM syndrome classification. B, representative tongue coating appearances. C, PCA of the four groups. D, unsupervised hierarchical clustering heatmap. E, KEGG pathway enrichment analysis for each TCM syndrome.

molecules (Figure 3C). The DNB members exhibited marked volatility and strong correlations with one another, indicating consistency of the critical stage across multiple liver diseases. Four groups of DNBs under different conditions were designated as DNB_{multi} (162 DNB genes), DNB_{CHB} (214 DNB genes), DNB_{LC} (22 DNB genes), and DNB_{HCC} (114 DNB genes).

KEGG pathway enrichment analysis of the four DNB groups (DNB_{multi}, DNB_{CHB}, DNB_{LC}, and DNB_{HCC}) revealed that the PI3K-AKT signaling pathway was consistently enriched and overlapped across all groups (Figure 4A – 4D). This shared pathway was therefore identified as a core molecular hub. Mapping of the DNB members onto PI3K-AKT signaling pathway showed their potential upstream regulatory roles (Figure 4E).

3.3 Dynamic changes of the DNB-DEG associated network before and after the critical transition

DNB characterized by strong correlations and marked volatility is recognized as a pivotal group of genes that signals the precursor of critical transitions during disease progression. Unlike DEGs, which represent specific dysfunctions in a static state, DNB captures dynamic transitions. A hypergeometric test was applied, and the results revealed no significant overlap between DNB and the 1 059 DEGs ($P > 0.05$).

Nevertheless, DNB is considered to occupy key positions, such as upstream locations, to coordinate or regulate DEGs involved in disease-associated functional processes. To further investigate the coordination between DNB members and DEGs during the critical transition from LGDHS to LKYDS, DNB members and related DEGs were merged as first-order neighbors within the whole human molecular network, thereby constructing the DNB-DEG associated network (97 nodes with 135 edges). At the pivotal transition, encompassing the period before and after LDSDS, significant dynamic changes were observed in the expression and co-regulation of this network (Figure 5A). Concurrently, functional analyses of the DNB-DEG association network revealed enrichment in various metabolic processes, consistent with previous observations of DNB alone. Notably, the PI3K-AKT signaling pathway was also enriched by members of the DNB-DEG association network, with DNB members identified upstream and DEGs downstream (Figure 5B). Consistently, GO biological process enrichment analysis (Figure 5C) demonstrated that the target biomarkers were predominantly enriched in metabolic process regulation, catabolic processes, and cell death/apoptotic processes.

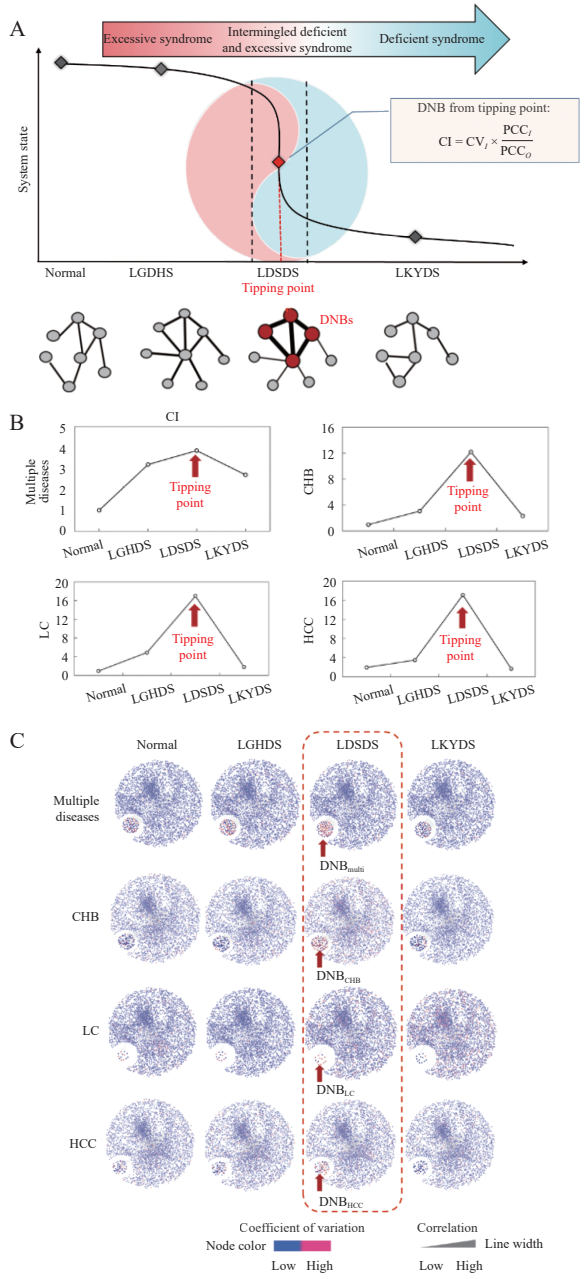

Figure 3 Identification of LDSDS as the critical tipping point using DNB analysis

A, tai-chi diagram illustrating DNB theory. Four simplified networks represent the normal state, LGDHS, LDSDS (critical tipping point), and LKYDS, respectively. Gray nodes indicate non-DNB molecules; red nodes indicate DNB. Node size reflects the CV of expression fluctuation, and edge thickness indicates the strength of Pearson correlation. B, CI for each disease group, with the tipping point at LDSDS. C, DNB network visualization.

3.4 Twenty-four core DNB members for predicting imminent TCM syndrome development

DNBs hold promises for early diagnosis of impending TCM syndrome development. Given the emphasis on the PI3K-AKT signaling pathway, 24 DNB members were selected as core DNBs, including *ITGB1*, *COL4A1*, *COL4A2*, *CCAAT*/enhancer-binding protein alpha (*CEBPA*), collagen type I alpha 1 chain (*COL1A1*), collagen type VI alpha 3 chain (*COL6A3*), cAMP responsive element binding protein 3-like 1 (*CREB3L1*), colony stimulating factor 1 (*CSF1*), *DDIT3*, ephrin A1 (*EFNA1*), epidermal growth factor receptor (*EGFR*), Fas ligand (*FASLG*), G protein subunit gamma 7 (*GNG12*), *GNG7*, cyclin-dependent kinase 4 (*CDK4*), interleukin 6 receptor (*IL6R*), interleukin 7 receptor (*IL7R*), interferon regulatory factor 5 (*IRF5*), insulin receptor substrate 1 (*IRS1*), integrin subunit alpha 11 (*ITGA11*), integrin subunit alpha 4 (*ITGA4*), Janus kinase 3 (*JAK3*), laminin subunit alpha 2 (*LAMA2*), and laminin subunit gamma 2 (*LAMC2*). These core DNBs include several transcription factors. Most of these DNB molecules were positioned upstream of the PI3K-AKT pathway (Figure 6), supporting the DNB theory that DNB may induce a transcriptional shift at the transcriptional level of downstream molecules.

The 24 core DNBs were analyzed by RF in the LGDHS, LDSDS, and LKYDS. The multidimensional scaling (MDS) plot (Figure 7A) showed clear separation among the four groups based on the 24 core DNB members. Next, the variable importance of candidate biomarkers was assessed using the Gini index derived from the random forest algorithm. A total of 13 candidate biomarkers were with Gini index > 1: *GNG12*, *GNG7*, *LAMA2*, *CSF1*, *ITGB1*, *CDK4*, *COL6A3*, *COL4A1*, *JAK3*, *COL4A2*, *DDIT3*, *EFNA1*, and *IRS1*. These prioritized genes were selected for subsequent functional validation and mechanistic investigation (Figure 7B).

To further select candidate genes from the 24 core DNB members for validation, ROC analysis was performed to evaluate their ability to distinguish LGDHS from LKYDS in CHB patients (Figure 8). This ROC analysis was independent of the RF analysis described above. Among the 24 genes, *ITGB1*, *COL4A1*, *COL4A2*, and *DDIT3* showed AUC > 0.70 (Figure 8A), indicating good discriminative performance. The remaining 20 genes exhibited lower AUC (< 0.70, Figure 8B). The four genes with AUC > 0.70 (*ITGB1*, *COL4A1*, *COL4A2*, and *DDIT3*) also ranked among the 13 important genes identified by RF analysis, and were therefore selected for subsequent independent validation.

According to DNB theory, the critical tipping point (LDSDS) represents a state of high molecular fluctuation characterized by an elevated CV and critical slowing-down. At this transient stage, direct comparison of DNB gene expression between pre- and post-transition states is of limited significance. Therefore, we focused validation of four core genes (*ITGB1*, *COL4A1*, *COL4A2*, and *DDIT3*) on the two states before and after the tipping point (LGDHS and LKYDS).

3.5 Expression validation of four core DNB genes in LGDHS and LKYDS across CLD

3.5.1 Comparison of core gene expression between LGDHS and LKYDS in the discovery dataset The expression levels of four core genes in the discovery dataset for LGDHS and LKYDS across multiple CLD diseases are

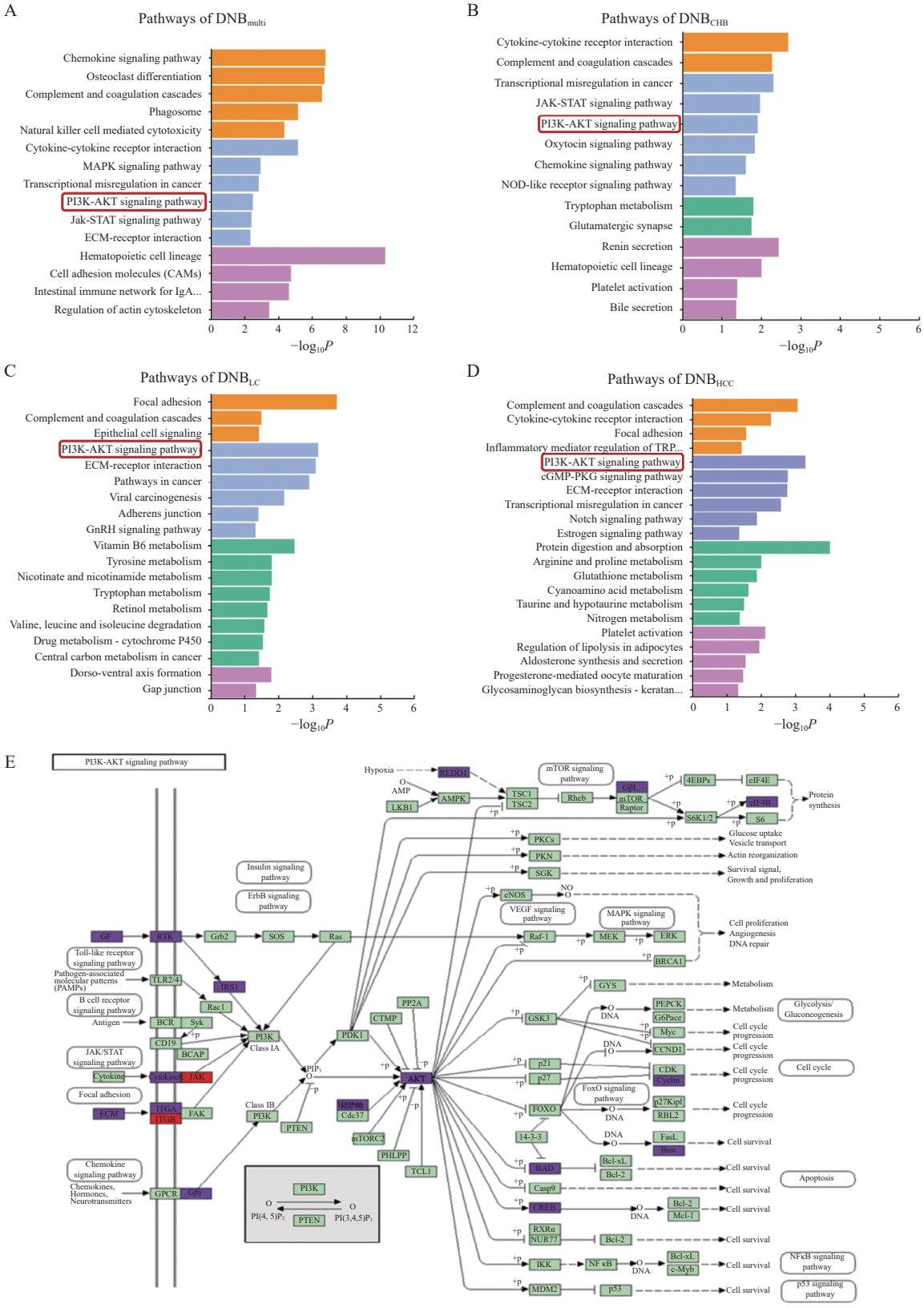

Figure 4 Overlapping KEGG pathways among all DNB groups and the resulting core pathways

A – D, KEGG pathways enriched in DNB_{multi}, DNB_{CHB}, DNB_{LC}, and DNB_{HCC}, respectively. The PI3K-AKT pathway, which overlapped across all four DNB groups, is highlighted in red. Cellular processes (orange); genetic information processing (light blue), metabolism (green), and organismal systems (purple). E, DNB members mapped onto the PI3K-AKT pathway.

shown in Figure 9A – 9D. All four genes were upregulated in LKYDS compared with LGDHS. Statistical analysis revealed significant differences for *ITGB1* (P = 0.040), *COL4A2* (P = 0.044), and *DDIT3* (P = 0.013), while *COL4A1* showed a non-significant upward trend (P = 0.088). Despite this borderline trend, *COL4A1* was

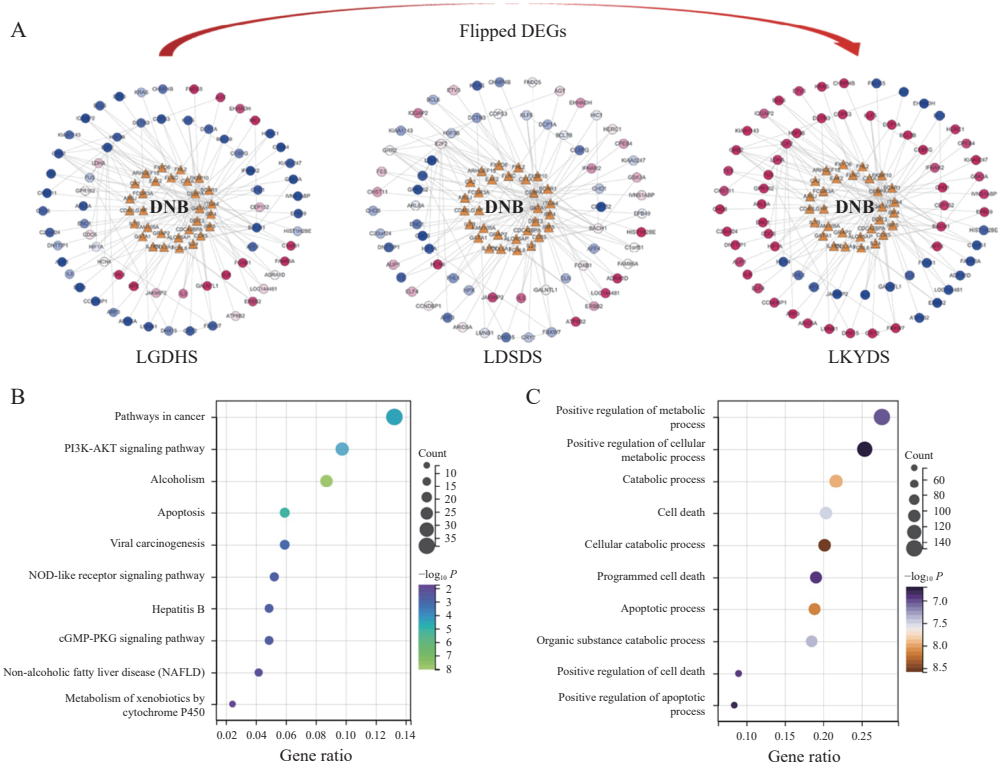

Figure 5 Functional characterization of the DNB-DEG association network

A, dynamic changes of the DNB-DEG network before and after the critical transition. B, KEGG pathway enrichment analysis of the DNB-DEG association network. C, GO biological process enrichment analysis of the DNB-DEG association network.

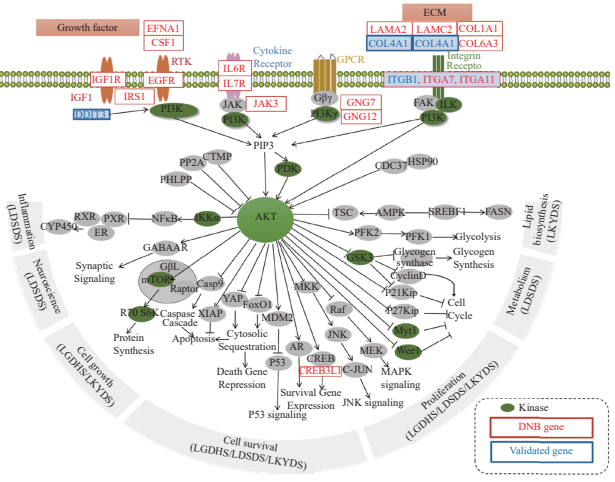

Figure 6 The 24 core DNB members localized upstream of the PI3K-AKT signaling pathway

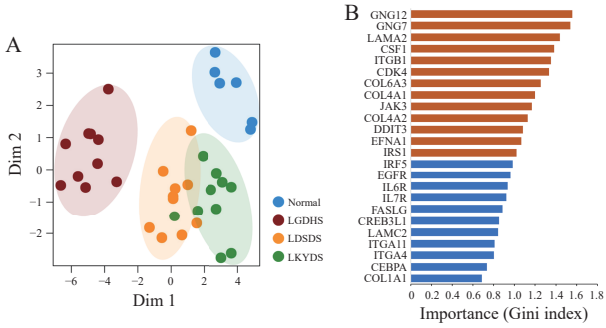

Figure 7 Classification of LGDHS, LDSDS, LKYDS, and normal control groups using core DNBs

A, MDS plot of the 24 core DNB members showing data distribution across the four groups. B, variable importance plot from RF analysis (performed on the three patient groups only) showing mean decrease in Gini impurity for each of the 24 core DNBs (orange: Gini > 1; blue: Gini < 1).

retained for further validation because it met three criteria: Gini score > 1, AUC = 0.707, and our gene selection was ROC-driven.

3.5.2 RT-qPCR validation in an independent LC cohort RT-qPCR assay was performed to validate four core DNB genes in an independent cohort of LC patients with LGDHS and LKYDS. At the transcript level, COL4A1 showed a slight increase in LKYDS compared with LGDHS, although this difference did not reach statistical significance (P = 0.120 1). All four genes were overexpressed in LKYDS relative to LGDHS (Figure 9E – 9H).

3.6 Disease progression and core gene expression patterns during chronic liver disease evolution

To determine whether the molecular features identified in TCM syndrome evolution also parallel disease severity progression, the expression patterns of the four core genes were examined in our discovery dataset, in the GSE89377 dataset from GEO, and in the TCGA-LIHC dataset across tumor stages (stage I, II, III, and IV). We validated the expression patterns of the four core genes across normal, CHB, LC, and HCC stages. In our internal dataset, all four genes were upregulated from normal to

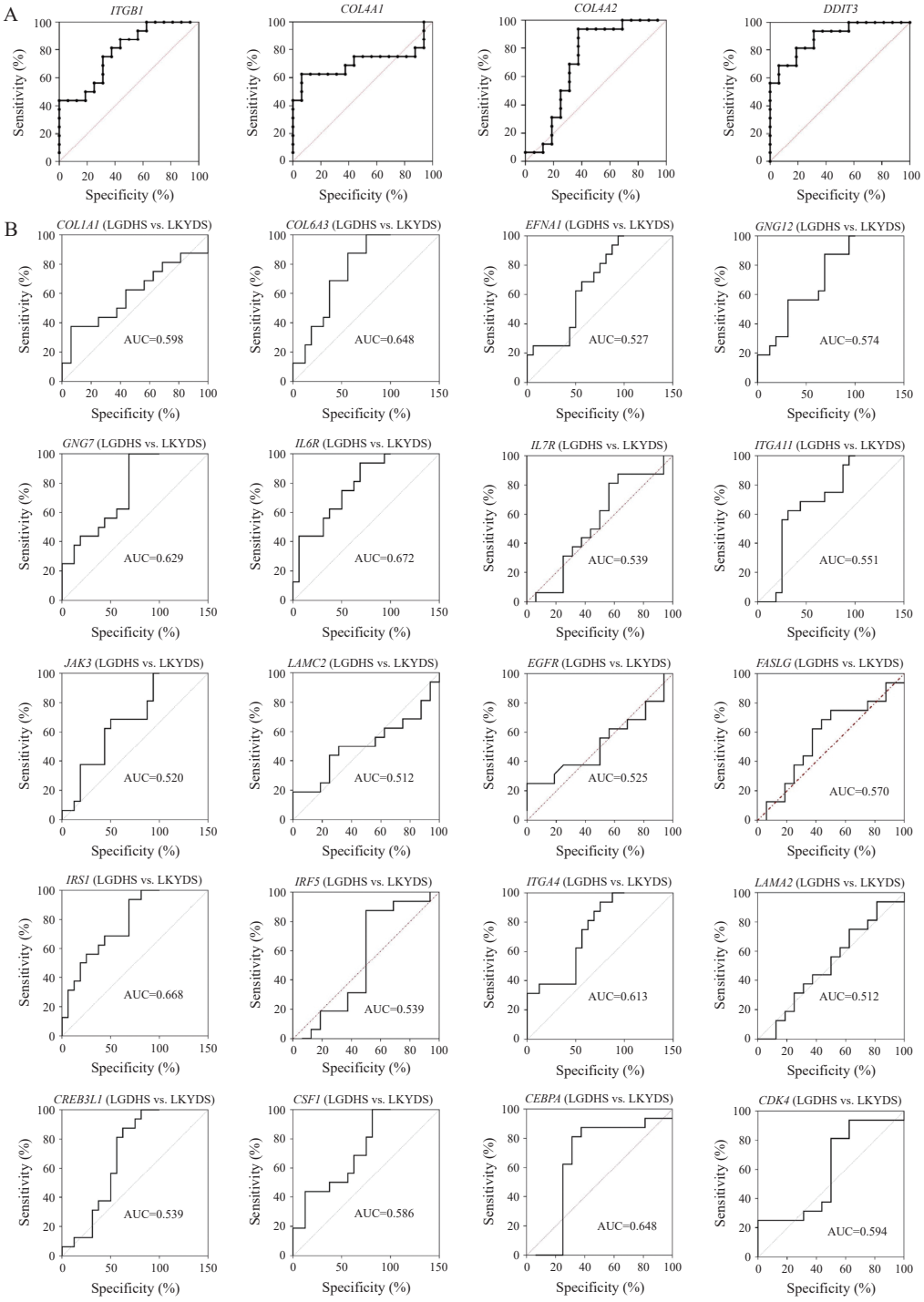

Figure 8 ROC analysis of 24 core DNB members for differentiating LGDHS from LKYDS in CHB patients

A, four genes with AUC > 0.70 (*ITGB1*, *COL4A1*, *COL4A2*, and *DDIT3*). B, the other 20 genes with AUC < 0.70 (*COL1A1*, *COL6A3*, *EFNA1*, *GNG12*, *GNG7*, *IL6R*, *IL7R*, *ITGA11*, *JAK3*, *LAMC2*, *EGFR*, *FASLG*, *IRS1*, *IRF5*, *ITGA4*, *LAMA2*, *CREB3L1*, *CSF1*, *CEBPA*, and *CDK4*).

HCC, with statistically significant differences observed by one-way ANOVA for *ITGB1* and *COL4A1* (Figure 10A – 10D). These findings were confirmed in the GSE89377 cohort, where the four genes exhibited significant stepwise increases during disease progression (Figure 10E – 10H). In the TCGA dataset, expression levels of the four genes progressively increased with advancing tumor stage (stage I – IV), further supporting their role in the malignant progression of chronic liver disease (Figure 10I – 10L). Multi-cohort TCGA data verified that these four genes (*ITGB1*, *COL4A1*, *COL4A2*, and *DDIT3*) are robust progressive markers during liver disease evolution. All four genes displayed consistent stepwise upregulation across liver cancer progression from stage I to IV. One-way ANOVA overall comparison revealed statistically significant differences among tumor stages for three genes (P = 0.019 for *ITGB1*, P = 0.038 for *COL4A1*, P = 0.021 for *COL4A2*), whereas *DDIT3* showed a consistent upward expression trend that did not reach statistical significance (P = 0.093).

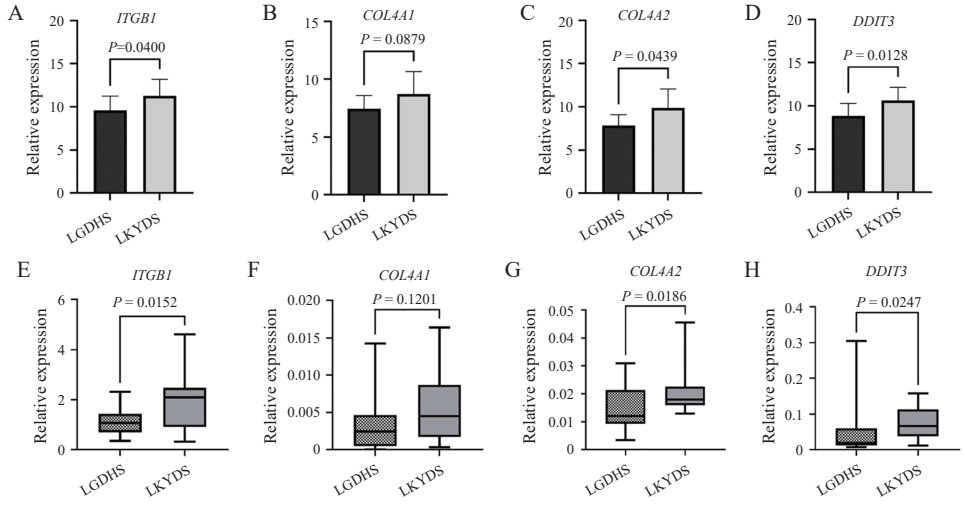

Figure 9 RT-qPCR validation of core DNBs for distinguishing LGDHS from LKYDS

A – D, expression levels of *ITGB1*, *COL4A1*, *COL4A2*, and *DDIT3* in the discovery cohort (microarray data) from patients with LGDHS (n = 11) and LKYDS (n = 9), respectively. E – H, validation of *ITGB1*, *COL4A1*, *COL4A2*, and *DDIT3* mRNA expression by RT-qPCR in an independent LC cohort (LGDHS, n = 30; LKYDS, n = 30), respectively.

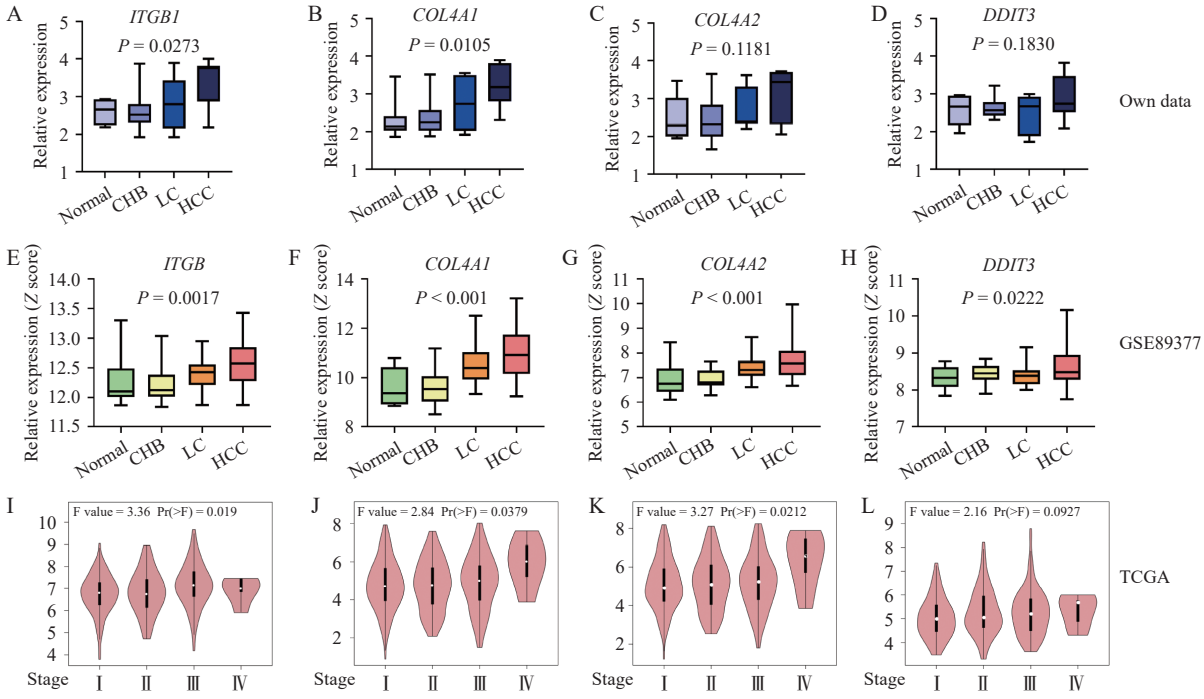

Figure 10 Expression patterns of four core genes during the progression from normal liver to HCC

A – D, expression levels of *ITGB1*, *COL4A1*, *COL4A2*, and *DDIT3* in the discovery cohort by disease stage, respectively (normal, n = 6; CHB, n = 12; LC, n = 10; HCC, n = 9). E – H, expression trends of *ITGB1*, *COL4A1*, *COL4A2*, and *DDIT3* in GSE89377, respectively (normal, n = 13; CHB, n = 20; LC, n = 34; HCC, n = 40). I – L, violin plots in TCGA-LIHC by tumor stage I – IV, respectively.

4 Discussion

4.1 LDSDS as the tipping point of TCM syndrome evolution and PI3K-AKT as the hub pathway

In this study, we applied DNB analysis to investigate the dynamic progression of TCM syndromes. We identified LDSDS as the critical tipping point in syndrome evolution, a finding consistently observed across all three diseases. The PI3K-AKT signaling pathway was enriched in DNB groups from every disease setting and was further identified as a core pathway in the DNB-DEG associated network, suggesting that it may serve as a molecular hub associated with the transition from excess (LGDHS) to deficiency (LKYDS).

The 24 core DNB members cover diverse functional categories, including ECM components [27], integrins [28], cytokine receptors [29], G proteins [30], receptor tyrosine kinases [31], transcription factors, and tumor necrosis factor (TNF) family genes [32]. Most of these DNB members are

positioned upstream of the PI3K-AKT pathway [31, 32], supporting the DNB concept that a dominant group of molecules exhibiting high fluctuation and strong correlation acts as a driver network triggering critical transitions. This upstream localization suggests that these DNB members may induce transcriptional changes in downstream molecules, thereby driving the shift from LGDHS to LKYDS via the LDSDS tipping point.

As anticipated, distinct dysfunctions in biological processes contribute to each typical TCM syndrome. According to TCM theory, metabolism plays a crucial role, as changes in TCM syndrome directly lead to the accumulation or overconsumption of metabolic waste and fluids [33]. Our results revealed that molecular signatures of LGDHS were enriched in immune system pathways (e.g., inflammatory mediator regulation of TRP channels) and the cytochrome P450-related pathway. CUI et al. [26] identified overexpression of nuclear factor kappa-light-chain-enhancer of activated B cells (NF-κB) and heat shock protein 70 (HSP70) in LGDHS, suggesting a connection to the NF-κB inflammatory pathway. Inflammatory genes such as high-sensitivity C-reactive protein (hs-CRP), IL-6, monocyte chemoattractant protein-1 (MCP-1), and cyclooxygenase-2 (COX-2) have also been reported to be upregulated in damp-heat TCM syndrome [27]. LGDHS refers to diseases and symptoms induced by an overabundanceofexogenouspathogenicfactorswhenthebody's resistance is not yet weakened. By contrast, LKYDS indicates diseases and symptoms caused by prolonged weakness and insufficiency of body resistance [20]. LDSDS was associated with dysfunctions in multiple metabolic pathways, particularly downstream pathways of the PI3K-AKT signaling pathway (e.g., the MAPK and JAK-STAT pathways). These results provide a molecular basis for the TCM concept of "same syndrome, different diseases", demonstrating that patients with different underlying liver diseases but the same TCM syndrome share common pathway dysregulations.

4.2 Diagnostic value of core DNB genes

The 24 core DNB members effectively identified the LDSDS transition. Notably, among these 24 DNBs, four genes—*ITGB1*, *COL4A1*, *COL4A2*, and *DDIT3*—exhibit high diagnostic ability. Integrins play a crucial role in cell-ECM attachment and signal transduction. Integrin members influence invasion by inducing epithelial-mesenchymal transition (EMT) in HCC, and their proangiogenic effect has also been reported [34]. Collagen IV is composed of an alpha 1 chain and an alpha 2 chain, encoded by *COL4A1* and *COL4A2*, respectively. Collagen IV is assembled into 112 heterotrimers and secreted into the ECM basement membrane [35]. *COL4A1* and *COL4A2* have been extensively studied concerning tumor biology. *COL4A1* is highly expressed in gastric cancer, colon cancer, and breast cancer, and is closely linked to the proliferation, differentiation, and migration of tumor cells [36]. WANG et al. [37] reported that FAK-Src signaling was a mechanism through which *COL4A1* overexpression facilitates the proliferation and metastasis of HCC cells. *DDIT3* is a transcription factor induced by multiple pathological conditions, including DNA damage, hypoxia, and amino acid starvation; its upregulated expression indicates DNA damage [38]. The mRNA expression patterns of these four genes support our hypothesis that LKYD represents a later phase than LGDHS.

4.3 Implications for TCM syndrome classification

The identification of LDSDS as a critical tipping point and the four core genes (*ITGB1*, *COL4A1*, *COL4A2*, and *DDIT3*) as potential quantitative markers may facilitate objective TCM syndrome differentiation, complementing traditional symptom-based diagnosis. Given the central role of the PI3K-AKT pathway across all three syndromes, targeted modulation of this pathway represents a potential strategy to block syndrome progression from excess to deficiency. In clinical practice, TCM formulas such as Xiaoyao San (Free and Easy Wanderer Powder) have been reported to regulate PI3K-AKT signaling [11], suggesting that early intervention at the LDSDS tipping point may help restore immune-metabolic balance and delay disease progression. Although direct causal links between *COL4A1*/*COL4A2*, *ITGB1*, and *DDIT3* have not been fully established, their concurrent upregulation from LGDHS to LKYDS suggests a coordinated functional network rather than a linear pathway. In CLD, ECM remodeling and ER stress are interconnected at the molecular level [39]. Thus, the four-gene signatures may reflect a multifaceted molecular response involving matrix reorganization, integrin signaling, and stress adaptation, jointly contributing to the transition from excess to deficiency in TCM syndrome evolution. Future studies are needed to dissect the hierarchical and temporal relationships among these genes.

4.4 Limitations and future prospects

However, several limitations should be noted. First, this study used cross-sectional samples from patients with TCM syndromes; longitudinal follow-up samples with dynamic syndrome changes are needed for further validation. Second, the transcriptomic discovery cohort was relatively small; although the findings were validated in independent cohorts, future studies with larger validation cohorts and incorporating cross-regional, cross-seasonal, or subgroup analyses would further strengthen the conclusions. Third, the ROC analysis was performed in the discovery CHB cohort; the diagnostic performance of the four-gene signature should be further evaluated in independent LC and HCC cohorts. Finally, mechanistic

experiments in animal models are required to establish causality between these genes and syndrome transition. In addition, clinical strategies such as "using drugs to validate syndromes" could be employed to further verify the association between these genes and TCM syndrome evolution.

5 Conclusion

This study identified LDSDS as the critical tipping point in TCM syndrome evolution across CHB, LC, and HCC, with the PI3K-AKT pathway consistently enriched as a core molecular hub. *ITGB1*, *COL4A1*, *COL4A2*, and *DDIT3* selected from the 24 core DNB members, were validated as potential quantitative markers for distinguishing excess state (LGDHS) from deficiency state (LKYDS). These findings provide a network-based scientific basis for the TCM concept of "same syndrome, different diseases" and offer candidate biomarkers and therapeutic targets for blocking syndrome progression in CLD.

Fundings

National Natural Science Foundation of China (82274183), and Special Project for the Development of Traditional Chinese Medicine in Xiamen (XWZY-2023-0615).

Ethical statement

This study was approved by the Ethics Committee of the Shanghai University of Traditional Chinese Medicine (2017-557-40-01), and written informed consent was obtained from all participants. The trial was retrospectively registered at ClinicalTrials.gov (NCT03189992).

Author contributions

Qingqing Chen: data curation, formal analysis, and writing – original draft. Hua Zhang: data collection. Dong Guo: data curation and formal analysis. Hong Cai: funding acquisition, project administration, and writing – review & editing. Yiyu Lu: funding acquisition, project administration, and writing – review & editing. All authors approved the submission and take responsibility for this manuscript.

Competing interests

The authors declare no conflicts of interest.

References

[1] WANG YC, HUANG YT, CHASE RC, et al. Global burden of digestive diseases: a systematic analysis of the global burden of diseases study, 1990 to 2019. Gastroenterology, 2023, 165(3): 773–783.e15.

[2] BALAKRISHNAN M, REHM J. A public health perspective on mitigating the global burden of chronic liver disease. Hepatology, 2024, 79(2): 451–459.

[3] MA LL, CHEN SZ, WANG HY, et al. Hepatitis B virus integration and hepatocarcinogenesis. Liver Research, 2025, 9(3): 189–198.

[4] LIU XL, LI MG, WANG XH, et al. Effects of adjuvant traditional Chinese medicine therapy on long-term survival in patients with hepatocellular carcinoma. Phytomedicine, 2019, 62: 152930.

[5] ZHENG Y, XIE L, YANG DJ, et al. Small-molecule natural plants for reversing liver fibrosis based on modulation of hepatic stellate cells activation: an update. Phytomedicine, 2023, 113: 154721.

[6] XUTIAN S, ZHANG J, LOUISE W. New exploration and understanding of traditional Chinese medicine. The American Journal of Chinese Medicine, 2009, 37(3): 411–426.

[7] LI NA, GONG YF, WANG J, et al. Integrating proteomics and targeted metabolomics to reveal the material basis of liver-gallbladder damp-heat syndrome in chronic hepatitis B. Digital Chinese Medicine, 2024, 7(4): 320–331.

[8] LU YY, FANG ZY, ZENG T, et al. Chronic hepatitis B: dynamic change in traditional Chinese medicine syndrome by dynamic network biomarkers. Chinese Medicine, 2019, 14(1): 52.

[9] LI H, WANG P, LI WJ, et al. Key issues and essential points in non-clinical pharmacology of traditional Chinese medicine formulas: enhancing the scientization and standardization. Acupuncture and Herbal Medicine, 2025, 5(4): 379–399.

[10] JIN XJ, WANG YR, WANG JR, et al. A machine learning approach to differentiate cold and hot syndrome in viral pneumonia integrating traditional Chinese medicine and modern medicine: machine learning model development and validation. JMIR Medical Informatics, 2025, 13: e64725.

[11] LU YY, LI MY, ZHOU QM, et al. Dynamic network biomarker analysis and system pharmacology methods to explore the therapeutic effects and targets of Xiaoyao San against liver cirrhosis. Journal of Ethnopharmacology, 2022, 294: 115324.

[12] HONG SX, JUN ZX, HUA SX, et al. Investigation of TCM syndrome distribution law in 594 patients with mild chronic hepatitis B. Chinese Journal of Integrated Traditional and Western Medicine on Liver Diseases, 2012, (3): 133–135.

[13] CHEN LN, LIU R, LIU ZP, et al. Detecting early-warning signals for sudden deterioration of complex diseases by dynamical network biomarkers. Scientific Reports, 2012, 2: 342.

[14] LU YY, ZHAO CQ, WANG CB, et al. The effect and mechanism of Qingre Huashi Formula in the treatment of chronic hepatitis B with Gan-Dan-Shi-Re syndrome: an integrated transcriptomic and targeted metabolomic analysis. Journal of Ethnopharmacology, 2024, 319: 117092.

[15] Chinese Society of Hepatology, CMA. Guideline for the Prevention and Treatment of Chronic Hepatitis B (2015 Edition). Journal of Practical Hepatology, 2016, 19(3): 389–400.

[16] LI PY, JING SW, TIAN G, et al. Revealing the critical state and identifying individualized dynamic network biomarker for type 2 diabetes through advanced analysis methods on individual basis. Scientific Reports, 2025, 15: 3925.

[17] HOU JL, LAI W. The guideline of prevention and treatment for

chronic hepatitis B: a 2015 update. Chinese Journal of Hepatology, 2015, 23(12): 888–905.

[18] Medical Administration Bureau, National Health Commission of the People's Republic of China. Guidelines for diagnosis and treatment of primary liver cancer in China (2019 edition). Chinese Journal of Hepatology, 2020, 28(2): 112–128.

[19] Branch of Hepatobiliary Diseases Caocm. The standards of traditional Chinese medicine syndrome differentiation for viral hepatitis. Journal of Clinical Hepatology, 2017, 33(10): 1839-1846.

[20] SONG YN, CHEN J, CAI FF, et al. A metabolic mechanism analysis of Fuzheng Huayu Formula for improving liver cirrhosis with traditional Chinese medicine syndromes. Acta Pharmacologica Sinica, 2018, 39(6): 942–951.

[21] GERTLER R, ROSENBERG R, FUEHRER K, et al. Detection of circulating tumor cells in blood using an optimized density gradient centrifugation. Molecular Staging of Cancer. Berlin, Heidelberg: Springer, 2003: 149–155.

[22] LU YY, ZHAO Y, SONG YN, et al. Serum cytokine profiling analysis for Zheng differentiation in chronic hepatitis B. Chinese Medicine, 2015, 10(1): 24.

[23] BREIMAN L. Random Forests. Machine Learning, 2001, 45(1): 5–32.

[24] MANDREKAR JN. Receiver operating characteristic curve in diagnostic test assessment. Journal of Thoracic Oncology, 2010, 5(9): 1315–1316.

[25] JIANG WY. Therapeutic wisdom in traditional Chinese medicine: a perspective from modern science. Trends in Pharmacological Sciences, 2005, 26(11): 558–563.

[26] CUI NJ, HU l, LAO SX. Relationship between Piwei damp-heat syndrome with expressions of nuclear factor-κB mRNA and heat shock protein 70 mRNA in patients with chronic gastritis. Chinese Journal of Integrated Traditional and Western Medicine, 2010, 30(1): 18–21.

[27] WANG J, HU YT, ZHAO KP, et al. Comprehensive analysis of the expression of cell adhesion molecules genes in hepatocellular carcinoma and their prognosis, and biological significance. Frontiers in Bioscience-Landmark, 2024, 29(2): 76.

[28] LIU TT, GU YM, ZHANG YY, et al. Integrin α2 in the microenvironment and the tumor compartment of digestive (gastrointestinal) cancers: emerging regulators and therapeutic opportunities. Frontiers in Oncology, 2024, 14: 1439709.

[29] O'SULLIVAN LA, LIONGUE C, LEWIS RS, et al. Cytokine receptor signaling through the Jak–Stat–Socs pathway in disease. Molecular Immunology, 2007, 44(10): 2497–2506.

[30] THOMPSON MD, CHIDIAC P, JOSE PA, et al. Genetic variants of accessory proteins and G proteins in human genetic disease. Critical Reviews in Clinical Laboratory Sciences, 2025, 62(2): 113–134.

[31] SHAMSAN E, ALMEZGAGI M, GAMAH M, et al. The role of PI3K/AKT signaling pathway in attenuating liver fibrosis: a comprehensive review. Frontiers in Medicine, 2024, 11: 1389329.

[32] GLAVIANO A, FOO ASC, LAM HY, et al. PI3K/AKT/mTOR signaling transduction pathway and targeted therapies in cancer. Molecular Cancer, 2023, 22(1): 138.

[33] DUAN LF, LUO Y, SHANG CR, et al. Therapeutic potential of Dahuang Mudan Decoction for severe acute pancreatitis: targeting oxidative stress and MAPK pathways through network pharmacology and experimental verification1. Journal of Ethnopharmacology, 2026, 360: 121116.

[34] GUO WJ, GIANCOTTI FG. Integrin signalling during tumour progression. Nature Reviews Molecular Cell Biology, 2004, 5(10): 816–826.

[35] KUO DS, LABELLE-DUMAIS C, GOULD DB. COL4A1 and COL4A2 mutations and disease: insights into pathogenic mechanisms and potential therapeutic targets. Human Molecular Genetics, 2012, 21(R1): R97–R110.

[36] GAO XY, ZHONG SH, TONG Y, et al. Alteration and prognostic values of collagen gene expression in patients with gastric cancer under different treatments. Pathology-Research and Practice, 2020, 216(3): 152831.

[37] WANG T, JIN HJ, HU JY, et al. COL4A1 promotes the growth and metastasis of hepatocellular carcinoma cells by activating FAK-Src signaling. Journal of Experimental & Clinical Cancer Research, 2020, 39(1): 148.

[38] LIN H, LIU SF, GAO WD, et al. DDIT3 modulates cancer stemness in gastric cancer by directly regulating CEBPβ. Journal of Pharmacy and Pharmacology, 2020, 72(6): 807–815.

[39] WANG C, ZHANG F, CAO Y, et al. Etoposide induces apoptosis in activated human hepatic stellate cells via ER stress. Scientific Reports, 2016, 6: 34330.

(Editor-in-Charge Jie Deng)

慢性肝病中医证候演变的分子特征：动态网络生物标志物分析

陈清清[a], 张华[b], 郭东[a], 蔡虹[c*], 陆奕宇[a*]

a. 上海中医药大学交叉科学研究院, 上海 201203, 中国
b. 上海中医药大学附属曙光医院肝病研究所肝脏和肾脏疾病重点实验室 (教育部), 上海 201203, 中国
c. 厦门市中医院肝病科一区, 福建 厦门 361015, 中国

【摘要】目的 从“异病同证”视角阐释慢性乙型肝炎（CHB）、肝硬化（LC）和肝细胞癌（HCC）患者中医证候的生物学基础，为慢性肝病（CLD）的诊断与治疗提供补充策略。**方法** 为探究中医证候在 CLD 中的动态演变，本研究对肝胆湿热证（LGDHS）、肝郁脾虚证（LDSDS）或肝肾阴虚证（LKYDS）的 CHB、LC 或 HCC 患者的外周血单个核细胞进行转录组学分析，受试者于 2018 年 8 月 1 日至 2021 年 12 月 31 日期间招募自上海中医药大学附属曙光医院。采用随机方差模型（RVM）*F* 检验筛选差异表达基因（DEGs），并经错误发现率（FDR）校正。运用主成分分析（PCA）和无监督层次聚类对样本分组进行可视化。采用动态网络生物标志物（DNB）分析识别证候演变中的关键转折阶段，随后对 DNB 成员进行基因本体（GO）和京都基因与基因组百科全书（KEGG）通路富集分析，以阐明其功能角色及参与的通路。利用随机森林（RF）分析和受试者工作特征（ROC）曲线下面积（AUC）对候选基因的重要性进行排序。采用独立 CLD 队列的数据（GSE89377）以及癌症基因组图谱肝细胞癌（TCGA-LIHC）数据库中的 RNA-seq 数据进行外部验证。此外，在独立的 LC 患者队列中采用逆转录定量聚合酶链式反应（RT-qPCR）验证候选基因的表达水平。**结果** 本研究共纳入 132 例受试者。DNB 分析表明，LDSDS 阶段是 CHB、LC 和 HCC 中医证候演变的关键转折点。PI3K-AKT 信号通路在所有 3 种 CLD 的 DNB 分析中均富集，提示其可能参与中医证候的关键转变。从 PI3K-AKT 通路的 24 个核心 DNB 成员中，通过 RF（基尼系数 > 1）和 ROC 分析鉴定出 4 个基因：整合素亚基 β1（*ITGB1*）、IV 型胶原 α1 链（*COL4A1*）、IV 型胶原 α2 链（*COL4A2*）和 DNA 损伤诱导转录因子 3（*DDIT3*）。ROC 分析显示，上述 4 个基因在区分 CHB 患者中的 LGDHS 与 LKYDS 时具有较高的判别能力，*ITGB1* 的 AUC 为 0.7891，*COL4A1* 为 0.7070，*COL4A2* 为 0.7148，*DDIT3* 为 0.8945。在独立 CLD 队列（GSE89377）中，4 个基因从正常至 CHB、LC、HCC 均呈显著阶梯性上调（$P < 0.05$）。在 TCGA-LIHC 数据集中，其表达随肿瘤分期逐渐升高。在独立 LC 队列（30 例 LGDHS vs. 30 例 LKYDS）中进行 RT-qPCR 实验验证，与 LGDHS 相比，LKYDS 中 *ITGB1*、*COL4A2* 和 *DDIT3* 的表达显著上调（分别为 $P = 0.0152$、0.0186 和 0.0247），而 *COL4A1* 呈上升趋势，但不显著（$P = 0.1201$）。**结论** 本研究为理解 CLD 中医证候演变的分子特征提供了一种新方法。PI3K-AKT 通路及其 4 个基因（*ITGB1*、*COL4A1*、*COL4A2*、*DDIT3*）在从实证（LGDHS）经关键 LDSDS 阶段向虚证（LKYDS）转变的过程中发挥重要作用。这些发现为中医证候分型提供了潜在的定量生物标志物和治疗靶点，并可能有助于阻断 CLD 的证候进展。

【关键词】中医证候演变；慢性肝病；慢性乙型肝炎；肝硬化；肝细胞癌；动态网络生物标志物；随机森林；PI3K-AKT 信号通路