



Article

Transcriptomic Identification of Diagnostic Biomarkers for Alcohol-Associated Liver Cirrhosis: Integration of Population-Level Epidemiology with Multi-Cohort Transcriptomic Analysis

Hao Wang ¹, Wenzhang Ding ², Linjie Zhang ³, Muyang Xu ¹ and Jing Sui ^{3,4,*}

- ¹ School of Electronic and Information Engineering, Anhui Jianzhu University, Hefei 230601, China; wh518812@stu.ahjzu.edu.cn (H.W.); myxu@stu.ahjzu.edu.cn (M.X.)
² School of Clinical Medicine, Wannan Medical University, Wuhu 241002, China; 24101010298@stu.wnmc.edu.cn
³ School of Public Administration, Nanjing University of Information Science and Technology, Nanjing 210044, China; zhanglinjie@njust.edu.cn
⁴ School of Public Health, Southeast University, Nanjing 210009, China
* Correspondence: suijing@nuist.edu.cn

Abstract

Alcohol-associated liver cirrhosis (ALC) lacks aetiology-specific molecular diagnostic biomarkers. This study aims to quantify the association between alcohol and cirrhosis risk, and to identify transcriptomic diagnostic biomarkers and candidate therapeutics. Methods: Survey-weighted logistic regression was applied to 17,007 adults from NHANES (2017–2023) to quantify alcohol-cirrhosis associations. ALC transcriptomic data from four GEO datasets were analysed using weighted gene co-expression network analysis (WGCNA) and three parallel machine learning algorithms (LASSO, Random Forest, SVM-RFE). External validation was performed in an independent cohort of 93 samples. Candidate therapeutics were identified via drug signature database querying and validated by molecular docking. Heavy drinking conferred a 5.14-fold increased cirrhosis risk (95% CI: 2.60–10.20, $p < 0.001$). Transcriptomic analysis revealed global downregulation of long non-coding RNAs (with 91.7% of dysregulated lncRNAs being suppressed). A five-gene diagnostic signature (*IL1B*, *CCL3*, *LUM*, *SPP1*, *ITGA6*), specifically developed to distinguish ALC from histologically normal liver tissue, achieved an area under the receiver operating characteristic curve (AUC) of 0.824 in an external validation cohort. Immune infiltration analysis uncovered global contraction of macrophage-associated transcriptomic signatures across M0, M1, and M2 subtypes, inversely correlated with fibrotic hub gene upregulation. Fluvastatin and honokiol were identified as candidate therapeutic agents, with strong binding affinities to *IL1B* and *CCL3*, respectively. This study confirms a dose-dependent alcohol-cirrhosis association and establishes a five-gene diagnostic signature (distinguishing ALC from normal liver tissue) alongside candidate therapeutics, warranting prospective clinical validation. The identified tissue-derived signature and therapeutic candidates provide a foundation for future ALC-specific diagnostic and therapeutic strategies; their translation into a non-invasive (e.g., blood-based) assay will require dedicated validation in circulating samples.



Academic Editor: Clare Selden

Received: 22 May 2026

Revised: 16 June 2026

Accepted: 23 June 2026

Published: 26 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\) license](https://creativecommons.org/licenses/by/4.0/).

Keywords: alcohol-associated liver cirrhosis; drug repurposing; diagnostic biomarkers; hepatic fibrosis; Immune microenvironment; machine learning; weighted gene co-expression network analysis

1. Introduction

Alcohol-associated liver cirrhosis (ALC) represents the end-stage of alcohol-related liver disease, contributing disproportionately to global cirrhosis-related mortality [1,2]. Unlike other aetiologies such as viral hepatitis or metabolic dysfunction-associated steatotic liver disease (MASLD), ALC is driven by direct ethanol toxicity and gut-derived endotoxemia [3]. Despite its profound clinical burden, therapeutic options to reverse established fibrosis remain severely limited, and current diagnosis relies heavily on invasive liver biopsy or non-specific clinical indices [4]. Consequently, validated, ALC-specific molecular biomarkers—and ultimately the non-invasive assays that could be derived from them—are urgently needed. Non-invasive diagnostic strategies for ALC remain inadequate; while transient elastography and serum fibrosis indices are widely used, none are ALC-specific, and their accuracy in distinguishing ALC-related fibrosis from other aetiologies remains limited [5].

Recent advances in high-throughput sequencing highlight the regulatory roles of long non-coding RNAs (lncRNAs) and mRNAs in hepatic fibrogenesis [6]. In ALC specifically, lncRNAs have been shown to regulate hepatocyte injury, lipid metabolism, inflammation, and hepatic stellate cell (HSC) activation through complex interactions with miRNAs and mRNAs [6,7]; however, their precise regulatory networks in end-stage alcoholic cirrhosis remain poorly characterised. Yet, in ALC specifically, the contribution of lncRNA-mediated regulatory networks to disease-specific fibrotic and immune programmes remains poorly characterised, partly because most transcriptomic studies have focused on protein-coding genes alone. While WGCNA and machine learning algorithms have successfully identified biomarkers in other liver diseases [8,9], and non-invasive diagnostic tools for ALD have advanced considerably in recent years [5], a comprehensive ALC-specific lncRNA-mRNA co-expression network analysis combined with multi-algorithm feature selection remains lacking. Furthermore, lncRNAs have emerged as key regulators of hepatocyte injury, inflammation, and fibrogenesis in ALD [6,7], yet their precise regulatory networks in end-stage alcoholic cirrhosis—particularly in concert with immune remodelling—remain poorly characterised.

A recent dose–response meta-analysis confirmed that alcohol consumption accelerates cirrhosis morbidity and mortality in a dose-dependent manner, with mortality risk accelerating at higher consumption levels [10]. To address this critical gap, we employed a two-stage integrative framework. First, we quantified the population-level association between heavy alcohol consumption and cirrhosis risk using data from 17,007 adults in the National Health and Nutrition Examination Survey (NHANES). Second, we integrated human ALC transcriptomic datasets to construct co-expression networks, applied a multi-algorithm ML approach (LASSO, Random Forest, and SVM-RFE) to isolate a robust diagnostic gene signature, and validated it in an independent external cohort. Finally, computational drug repurposing and molecular docking were conducted to nominate novel therapeutic candidates. Together, this framework is designed to yield a tissue-based diagnostic signature and mechanistically grounded therapeutic leads for ALC, providing a rational basis for subsequent non-invasive assay development.

2. Results

2.1. Baseline Characteristics and the Epidemiological Link Between Alcohol and Cirrhosis

Among the 17,007 NHANES participants, 80 were identified with liver cirrhosis. As summarised in Table 1, patients with cirrhosis were significantly older (58.81 vs. 48.56 years, $p < 0.001$) and had a higher prevalence of diabetes and smoking (both $p < 0.001$). Clinically, the cirrhosis cohort exhibited markedly worse hepatic profiles, including elevated AST, ALT, and FIB-4 index, alongside reduced platelet counts (all $p < 0.001$). Notably, the proportion of heavy drinkers was more than three-fold higher in the cirrhosis group than in the non-cirrhosis group (49.8% vs. 14.7%, $p < 0.001$).

Table 1. Baseline characteristics of study participants by liver cirrhosis status, NHANES 2017–2023.

Characteristics	Overall	Non-Cirrhosis	Cirrhosis	<i>p</i> Value
Participants, <i>n</i>	17,007	16,927	80	
Age, years, mean (SD)	48.60 (17.44)	48.56 (17.44)	58.81 (10.48)	<0.001
Gender, <i>n</i> (%)				0.318
Male	8393 (48.3)	8349 (48.3)	44 (55.5)	
Female	8614 (51.7)	8578 (51.7)	36 (44.5)	
Race/Ethnicity, <i>n</i> (%)				0.470
Mexican American	3285 (7.7)	3273 (7.7)	12 (12.3)	
Other Hispanic	2361 (8.6)	2351 (8.6)	10 (9.2)	
Non-Hispanic White	5752 (61.4)	5718 (61.3)	34 (59.1)	
Non-Hispanic Black	3646 (11.3)	3636 (11.3)	10 (5.3)	
Other/Multi-racial	1963 (11.1)	1949 (11.1)	14 (14.0)	
Education level, <i>n</i> (%)				0.099
Below high school	2916 (10.5)	2898 (10.5)	18 (18.6)	
High school/GED	4008 (26.7)	3986 (26.7)	22 (33.5)	
Some college	4882 (30.0)	4860 (29.9)	22 (33.6)	
College or above	5154 (32.8)	5136 (32.8)	18 (14.3)	
Family income-to-poverty ratio, <i>n</i> (%)				<0.001
Below poverty	2748 (13.0)	2732 (13.0)	16 (28.4)	
Near poverty	5416 (35.2)	5385 (35.2)	31 (54.0)	
Above poverty	6515 (51.8)	6504 (51.8)	11 (17.6)	
BMI, kg/m ² , mean (SD)	29.74 (7.32)	29.73 (7.32)	31.03 (7.09)	0.273
BMI category, <i>n</i> (%)				0.385
<25	4189 (27.0)	4175 (27.1)	14 (16.8)	
25–<30	5142 (31.9)	5120 (31.9)	22 (34.3)	
≥30	5349 (41.1)	5305 (41.1)	44 (48.9)	
Sleep duration, hours, mean (SD)	7.64 (1.52)	7.64 (1.52)	7.72 (2.32)	0.870
Diabetes status, <i>n</i> (%)				<0.001
No	14,223 (85.7)	14,163 (85.7)	60 (62.9)	
Borderline	553 (2.7)	546 (2.7)	7 (7.9)	
Yes	2216 (11.6)	2203 (11.6)	13 (29.1)	
Smoking status, <i>n</i> (%)				<0.001
Never/Unknown	9997 (60.0)	9963 (60.0)	34 (30.7)	
Former	4274 (24.4)	4243 (24.4)	31 (49.1)	
Current	2736 (15.6)	2721 (15.6)	15 (20.2)	
Physical activity, <i>n</i> (%)				0.062
Active	5403 (54.6)	5378 (54.6)	25 (72.7)	
Inactive	4371 (45.4)	4355 (45.4)	16 (27.3)	
Alcohol drinking status, <i>n</i> (%)				<0.001
Never	1199 (7.7)	1195 (7.7)	4 (5.0)	
Non-heavy	9924 (77.6)	9899 (77.6)	25 (45.2)	
Heavy	2022 (14.7)	1989 (14.7)	33 (49.8)	
Ever heavy drinker, <i>n</i> (%)				<0.001
No	9937 (84.1)	9899 (84.1)	38 (48.0)	
Yes	2014 (15.9)	1984 (15.9)	30 (52.0)	
Avg drinks/day, median [IQR]	2.00 [1.00, 3.00]	2.00 [1.00, 3.00]	2.05 [2.00, 4.00]	0.007
AST, U/L, median [IQR]	20.00 [16.00, 24.00]	20.00 [16.00, 24.00]	29.11 [22.00, 58.10]	<0.001
ALT, U/L, median [IQR]	18.00 [13.00, 26.00]	18.00 [13.00, 26.00]	25.00 [20.00, 44.01]	<0.001
Platelet count, 10 ⁹ /L, median [IQR]	245.00 [208.00, 288.00]	245.00 [208.00, 288.00]	183.21 [117.21, 250.68]	<0.001
FIB-4 index, median [IQR]	0.86 [0.57, 1.30]	0.86 [0.57, 1.30]	2.27 [1.36, 3.28]	<0.001
NHANES Cycle, <i>n</i> (%)				0.233
2017–2020	9213 (49.6)	9179 (49.6)	34 (39.3)	
2021–2023	7794 (50.4)	7748 (50.4)	46 (60.7)	

Data are presented as weighted mean (SD) for normally distributed continuous variables, median [IQR] for skewed continuous variables, or unweighted *n* (weighted %) for categorical variables. *p* values were calculated using a survey-weighted *t*-test for continuous variables and the Rao–Scott chi-squared test for categorical variables. Abbreviations: BMI, body mass index; IQR, interquartile range; AST, aspartate aminotransferase; ALT, alanine aminotransferase; FIB-4, Fibrosis-4 index; GED, General Educational Development; NHANES, National Health and Nutrition Examination Survey.

To determine whether these group differences reflected an independent association, multivariable, survey-weighted logistic regression was performed (Table 2). In the fully adjusted model (Model 3), which accounted for demographics, education, BMI, and diabetes, a history of heavy drinking independently conferred a 5.14-fold increased risk of cirrhosis (95% CI: 2.60–10.20, $p < 0.001$). Furthermore, a significant dose–response relationship was observed: each additional daily alcoholic drink increased the risk of cirrhosis by 18% (OR = 1.18, 95% CI: 1.07–1.30, $p = 0.002$). These population-level findings establish a robust epidemiological foundation for the subsequent molecular analyses.

Table 2. Survey-weighted logistic regression analysis of the association between alcohol consumption and liver cirrhosis, NHANES 2017–2023.

Exposure	N	Events	Model 1 (Crude)	Model 2	Model 3
Analysis A: Ever heavy drinker					
No (reference)	9937	38	1.00	1.00	1.00
Yes	2014	30	5.73 (2.81–11.70) ***	5.48 (2.85–10.60) ***	5.14 (2.60–10.20) ***
<i>p</i> for trend			<0.001	<0.001	<0.001
Analysis B: Drinking status					
Never (reference)	1199	4	1.00	1.00	1.00
Non-heavy	9924	25	0.89 (0.24–3.30)	1.12 (0.30–4.11)	1.17 (0.32–4.30)
Heavy	2022	33	5.20 (1.53–17.70) *	6.30 (1.91–20.80) **	6.15 (1.87–20.20) **
<i>p</i> for trend			<0.001	<0.001	<0.001
Analysis C: Average drinks per day (continuous)					
Per 1-drink/day increase	9522	35	1.15 (1.05–1.26) **	1.19 (1.08–1.32) **	1.18 (1.07–1.30) **
<i>p</i> for trend			0.005	0.002	0.002

Values are presented as OR (95% CI). All analyses incorporated NHANES survey weights to account for the complex sampling design. Model 1: Unadjusted. Model 2: Adjusted for age, gender, and race/ethnicity. Model 3: Further adjusted for education, BMI, and diabetes status. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Heavy drinking was defined based on NHANES questionnaire ALQ151 (“Ever have 4/5 or more drinks every day?”). Abbreviations: OR, odds ratio; CI, confidence interval; BMI, body mass index; NHANES, National Health and Nutrition Examination Survey.

2.2. Transcriptomic Landscape and Global lncRNA Suppression

Differential expression analysis of ALC tissues identified 1471 differentially expressed genes (DEGs; adjusted $p < 0.05$ and $|\log_2FC| > 1$) (Figure 1). While protein-coding mRNAs exhibited a relatively balanced pattern of dysregulation, long non-coding RNAs (lncRNAs) displayed a profoundly asymmetric pattern: 332 were significantly downregulated versus only 30 upregulated. This pronounced asymmetry (91.7% downregulated) suggests that global lncRNA silencing may represent a candidate molecular feature of ALC pathophysiology, warranting further investigation. Functional enrichment revealed that up-regulated DEGs were predominantly involved in extracellular matrix (ECM) organisation and PI3K-Akt signalling, whereas downregulated genes were enriched in immune-related signalling, ion homeostasis, and xenobiotic metabolism pathways (Figure 2). To reduce visual complexity, only the top 10 most significant pathways (ranked by adjusted p -value) are displayed for each comparison.

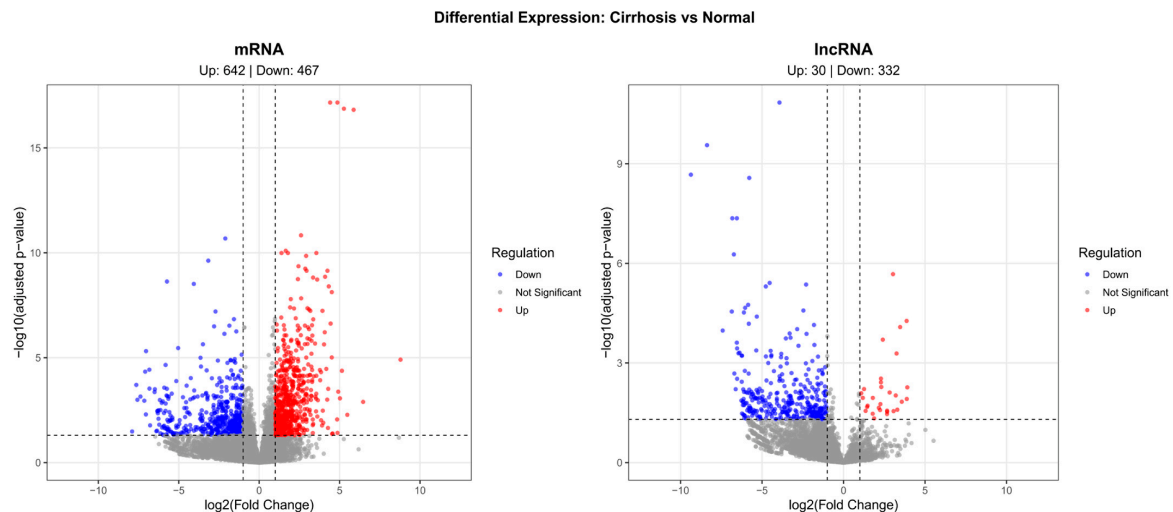


Figure 1. Differential expression of mRNAs and lncRNAs in alcoholic liver cirrhosis. Volcano plots of differentially expressed genes between ALC ($n = 6$) and normal liver tissues ($n = 12$). (A) mRNA: 642 upregulated (red), 467 downregulated (blue). (B) lncRNA: 30 upregulated (red), 332 downregulated (blue). Grey: not significant. Dashed lines: adjusted p -value = 0.05 and $|\log_2FC| = 1$.

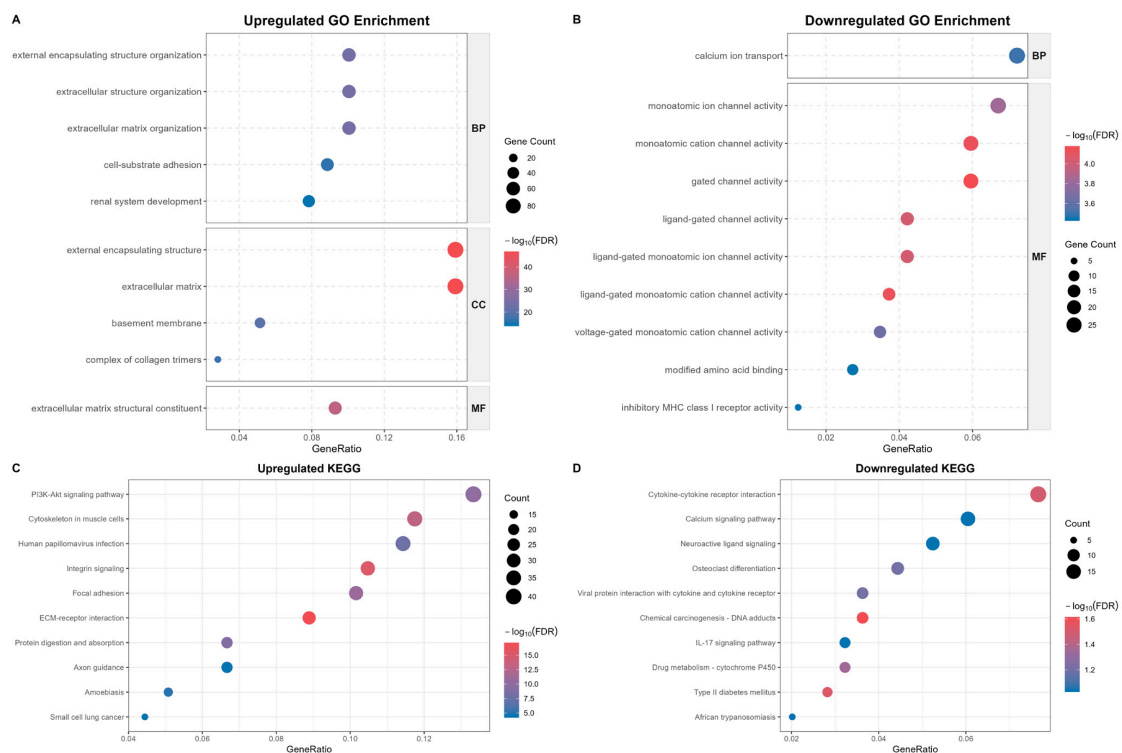


Figure 2. Functional enrichment analysis of DEGs. (A,B) Top 10 GO terms for upregulated (A) and downregulated (B) genes, faceted by biological process (BP), cellular component (CC), and molecular function (MF). (C,D) Top 10 KEGG pathways for upregulated (C) and downregulated (D) genes. Pathways were ranked by adjusted p -value (FDR), with only the globally most significant 10 terms shown per direction to enhance visual clarity. Dot size represents gene count, and colour represents $-\log_{10}(\text{FDR})$.

2.3. Weighted Gene Co-Expression and Interaction Networks

To capture system-level properties, WGCNA was applied, identifying nine distinct co-expression modules (Figure 3). Module-trait correlation analysis highlighted the “Pink” module as having the strongest negative correlation with ALC, while the “Yellow” module

showed the most significant positive correlation. Consistent with overall ALC pathology, the Yellow module was heavily enriched in fibrogenesis and ECM–receptor interactions, whereas the Pink module was enriched in immune signalling pathways (Figure 4). The most significantly enriched terms (top 10 by adjusted *p*-value) for each module are shown in the figure. The gene composition of all nine modules is summarised in Supplementary Table S1.

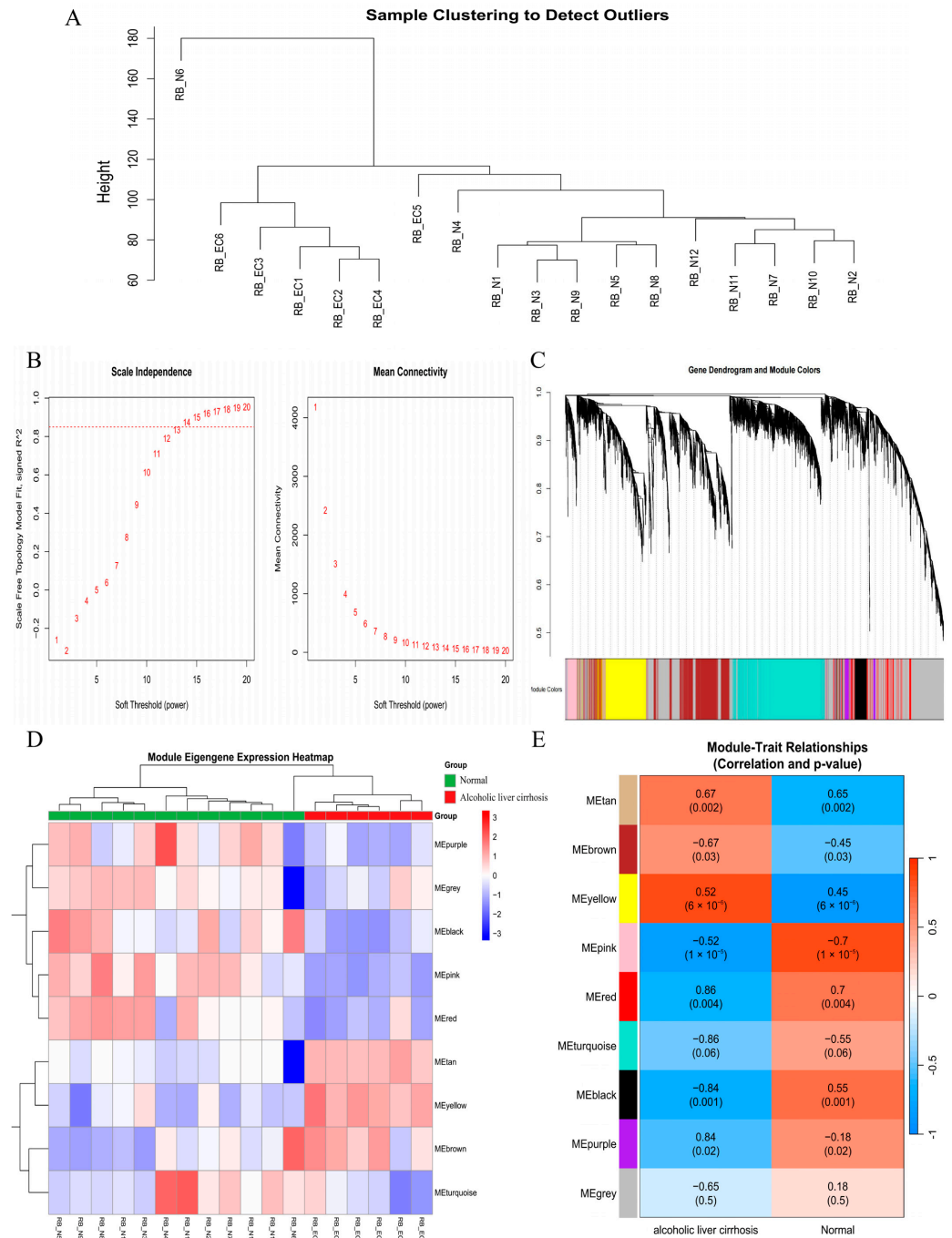


Figure 3. Construction of the weighted gene co-expression network. (A) Sample clustering dendrogram. (B) Scale-free topology fit index (signed R^2 , left) and mean connectivity (right) across candidate soft-thresholding powers. In the left panel, each number denotes a candidate soft-thresholding power, and the red dashed line marks the scale-free topology fit threshold (signed $R^2 = 0.85$); the smallest power reaching this threshold ($\beta = 14$) was selected for network construction. (C) Gene clustering dendrogram with 9 identified modules. (D) Module eigengene expression heatmap. (E) Module–trait relationship heatmap.

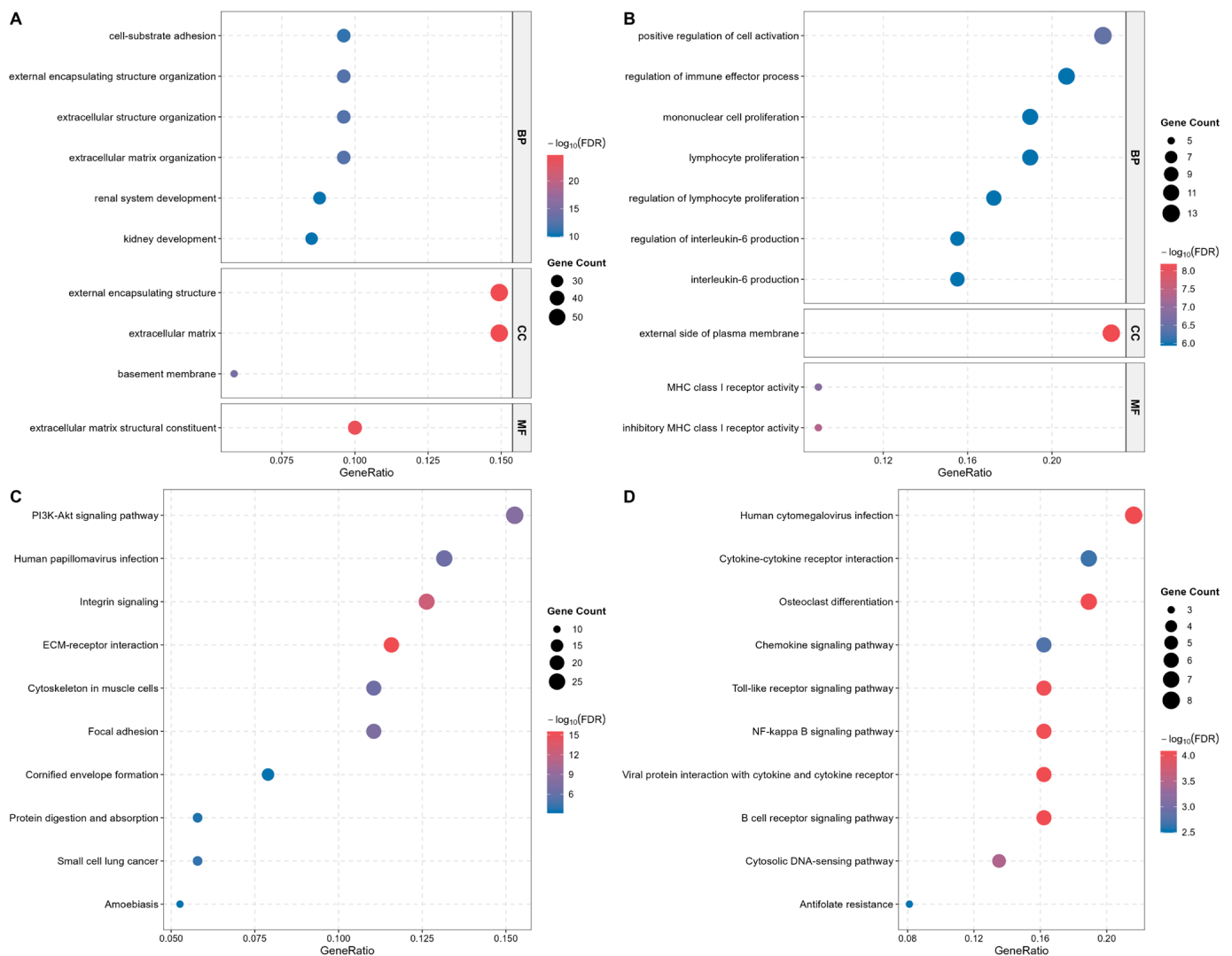


Figure 4. Functional enrichment analysis of the Yellow and Pink modules. (A,C) Top 10 GO terms (A) and KEGG pathways (C) for the Yellow module, which was enriched in fibrosis-related functions. (B,D) Top 10 GO terms (B) and KEGG pathways (D) for the Pink module, which was enriched in immune-related functions. Pathways were ranked by adjusted p -value (FDR), with only the 10 most significant terms shown per module to reduce complexity. Dot size reflects gene count; colour gradient represents $-\log_{10}(\text{FDR})$.

Comprehensive lncRNA-mRNA co-expression and pathway interaction networks for the non-key modules are provided in Supplementary Figures S1 and S2. For the critical ALC-associated modules, the Pink network centred around the downregulated lncRNA TMEM26-AS1, linking to inflammatory mediators, whereas the Yellow network was orchestrated by the upregulated lncRNA UCA1, driving fibrotic pathways (Supplementary Figure S3). Subsequent protein-protein interaction (PPI) network construction prioritised core regulatory nodes corresponding to these modules (Figure 5).

2.4. Machine Learning-Based Diagnostic Signature Identification

To derive a candidate diagnostic panel from these hub genes, three parallel machine learning algorithms—LASSO regression, Random Forest, and SVM-RFE—were applied. The detailed feature selection processes and algorithmic metrics for the Pink and Yellow modules are systematically documented in Supplementary Figures S4 and S5, respectively. The intersection of these algorithms identified a five-gene core signature: *IL1B*, *CCL3*, *LUM*, *SPP1*, and *ITGA6*. A diagnostic model combining these five features achieved a

corrected Area Under the Curve (AUC) of 0.922 in the training cohort when evaluated via leave-one-out cross-validation (LOOCV) (Supplementary Figure S6). Given the limited training sample size ($n = 18$), the training AUC (0.922) should be interpreted cautiously; the externally validated AUC of 0.824 in 93 independent samples is considered the primary performance estimate.

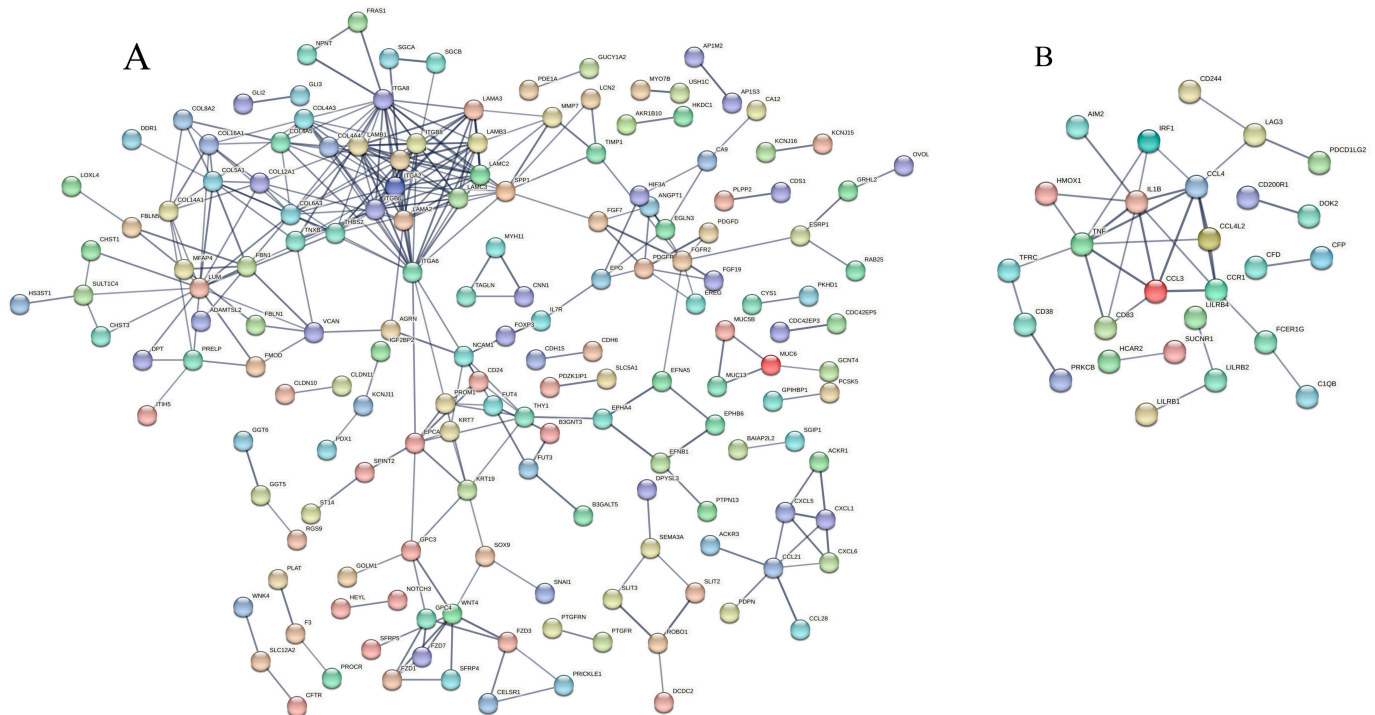


Figure 5. PPI networks of the Pink and Yellow modules. (A) Pink module: 57 nodes, 34 edges. (B) Yellow module: 389 nodes, 310 edges. Networks constructed using STRING database (interaction score ≥ 0.700).

2.5. CTD-Based Validation of Ethanol–Gene Interactions

To confirm that the five-gene diagnostic signature is biologically responsive to ethanol exposure, we systematically queried the Comparative Toxicogenomics Database (CTD). Direct ethanol–gene interaction evidence was identified for four of the five genes, each supported by distinct peer-reviewed studies. Specifically, CTD records show that ethanol exposure upregulates *IL1B* [11–15] and *CCL3* [14,16,17] and increases *SPP1* [18,19] and *ITGA6* [20]. For *LUM*, the CTD does not record a direct ethanol–gene interaction. Instead, an indirect association was identified: combined exposure to carbon tetrachloride and ethanol has been reported to upregulate *LUM* mRNA expression [21]. While this indirect evidence aligns with *LUM*'s established role in alcohol-driven fibrogenesis, it should be interpreted with caution, as the co-exposure context precludes the attribution of this transcriptional alteration to ethanol alone. Collectively, direct ethanol-modulated expression evidence was identified for four of the five signature genes (*IL1B*, *CCL3*, *SPP1*, *ITGA6*), with *LUM* supported by indirect co-exposure data. This nuanced pattern reinforces the biological plausibility of our diagnostic signature while maintaining stringency regarding chemical–gene causal interactions.

2.6. External Validation and Characterisation of Macrophage-Associated Transcriptomic Signatures

The diagnostic performance of the five-gene signature was further confirmed in the independent validation cohort, with individual gene expression trajectories remaining consistent across all three contributing datasets (Figure 6).

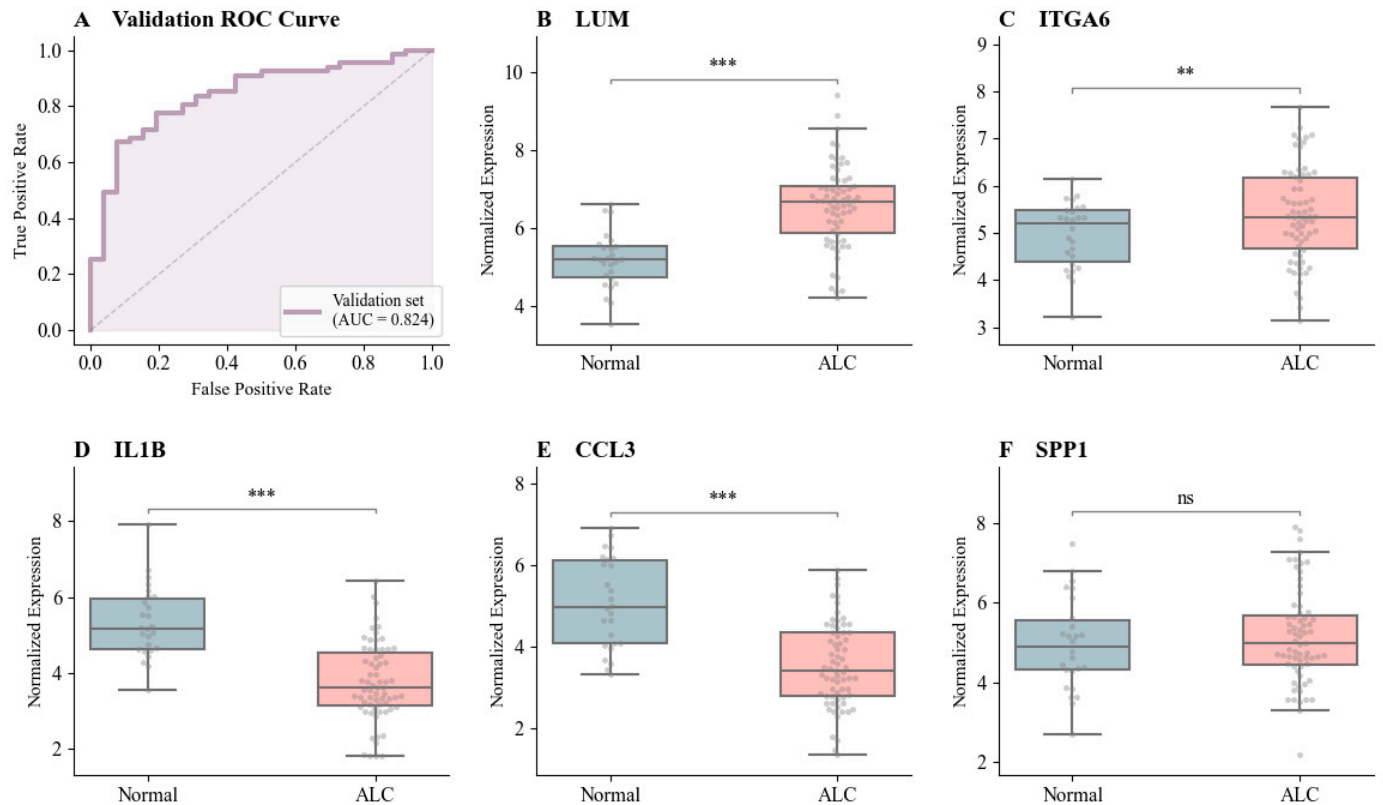


Figure 6. External validation of the diagnostic model and core genes in an independent cohort ($n = 93$). (A) ROC curve of the 5-gene diagnostic model in the external validation cohort. (B–F) Boxplots displaying the expression levels of core genes (*LUM*, *ITGA6*, *IL1B*, *CCL3*, *SPP1*) in normal (blue) vs. ALC (pink) tissues. Statistical significance was determined by Wilcoxon rank-sum test (** $p < 0.01$, *** $p < 0.001$, ns: not significant). AUC = 0.824.

To unravel the immunological context of these alterations, ssGSEA was performed. Strikingly, ALC tissues exhibited a systemic contraction of the macrophage pool—including M0, M1, and M2 subtypes. This attenuation of macrophage-associated transcriptomic signatures was inversely correlated with the upregulation of fibrotic hub genes (*LUM*, *SPP1*, *ITGA6*) (Figure 7). This co-occurrence is consistent with—but does not by itself demonstrate—a model in which extensive ECM deposition accompanies a reduced representation of macrophage-associated transcripts. Whether this reflects physical exclusion of immune effectors or an immunosuppressive niche remains a hypothesis to be tested by spatially resolved or single-cell approaches.

2.7. Computational Drug Repurposing and Molecular Docking

Leveraging the 5-gene signature, a query of the DSigDB identified several compounds predicted to reverse the ALC-associated transcriptomic pathology, with fluvastatin and honokiol emerging as top candidates. The top-ranked compounds identified from DSigDB querying are summarised in Supplementary Table S2. Molecular docking simulations confirmed stable binding interactions, with fluvastatin achieving a binding affinity of -7.65 kcal/mol to *IL1B* and honokiol of -6.94 kcal/mol to *CCL3* (Figure 8). Specifically,

fluvastatin selectively engaged the active domain of IL1B through critical hydrogen bonds, while honokiol demonstrated preferential binding to CCL3.

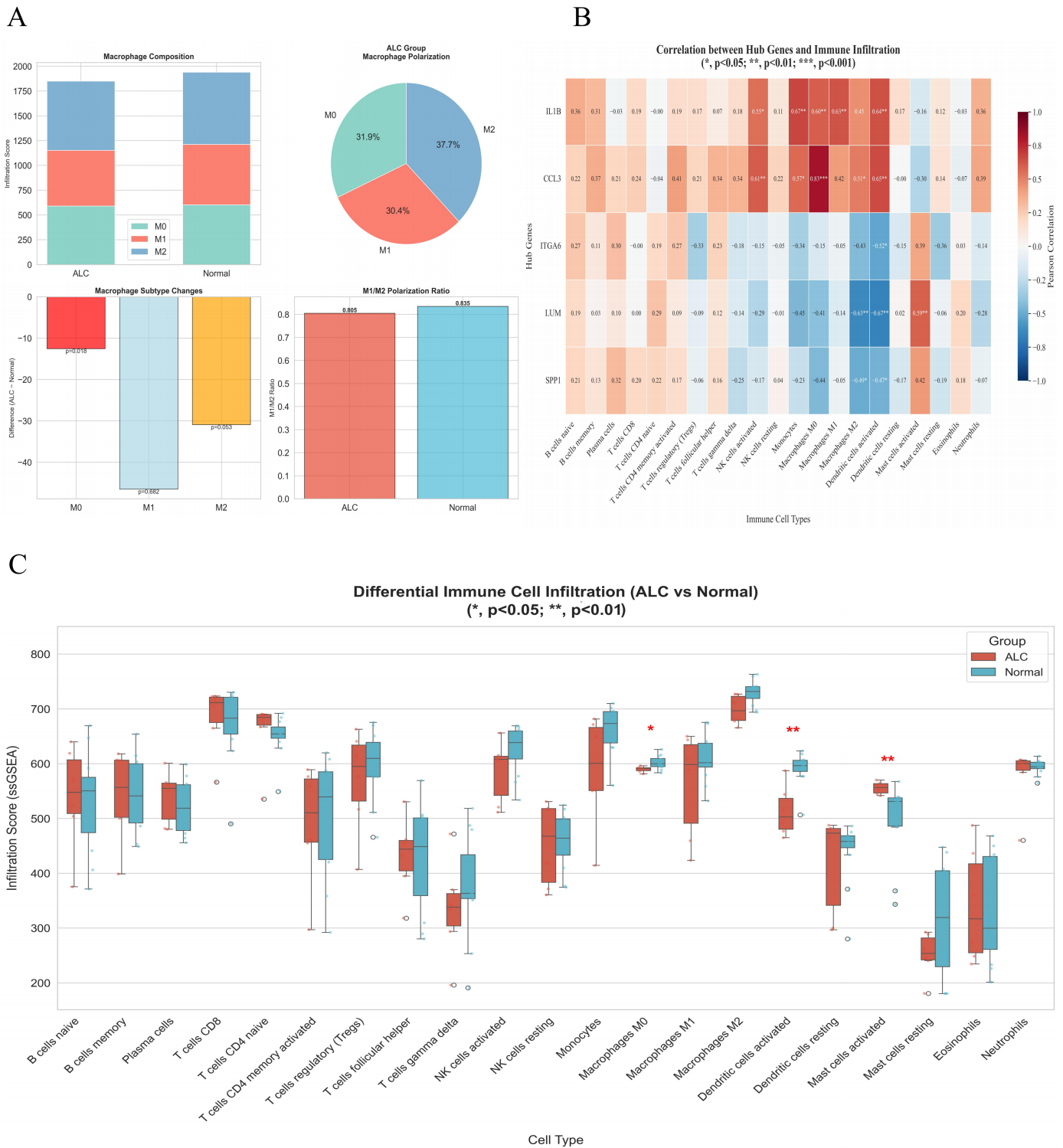


Figure 7. Immune infiltration landscape and hub gene-immune correlations in ALC. (A) Macrophage dynamics showing global attenuation of macrophage-associated transcriptomic signatures and M1/M2/M0 distribution. (B) Hub gene-immune correlation heatmap for inflammatory hubs (*IL1B*, *CCL3*) and fibrotic hubs (*LUM*, *SPPI*, *ITGA6*). (C) Differential infiltration via ssGSEA scores.

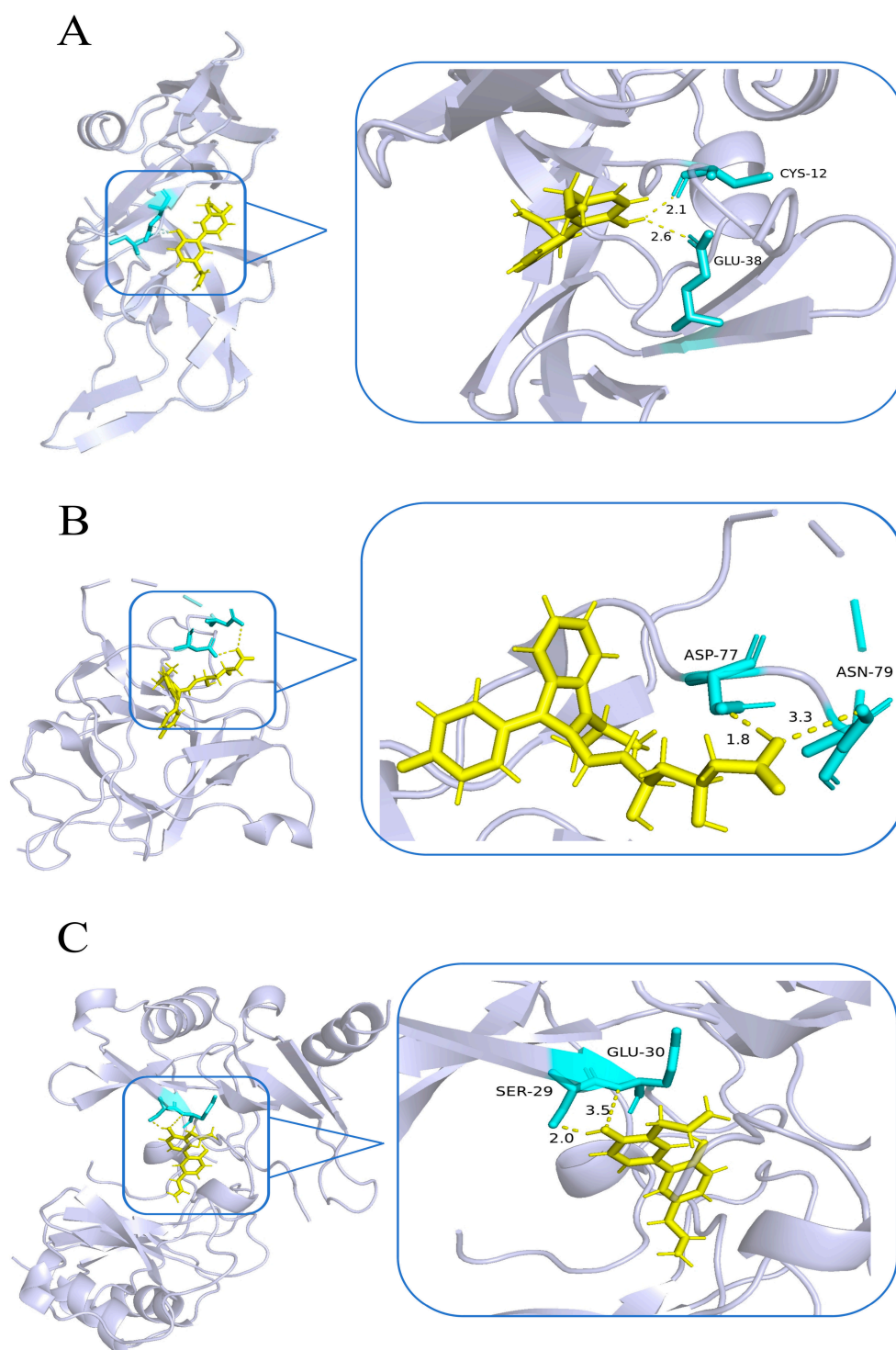


Figure 8. Molecular docking analysis of candidate compounds with target proteins. Three-dimensional visualisation of optimal binding poses. The protein backbone is shown in cartoon representation (light purple), ligands in yellow sticks, and key residues in cyan. Hydrogen bonds are indicated by yellow dashed lines, with distances in angstroms. (A) Honokiol-CCL3 (3FPR): hydrogen bonds with CYS-12 (2.1 Å) and GLU-38 (2.6 Å). (B) Fluvastatin-IL1B (3POK): hydrogen bonds with ASP-77 (1.8 Å) and ASN-79 (3.3 Å). (C) Honokiol-CCL3 (4MHE): hydrogen bonds with SER-29 (2.0 Å) and GLU-30 (3.5 Å).

3. Discussion

This integrative study advances the molecular understanding of ALC through two complementary lines of evidence. At the population level, heavy alcohol consumption conferred a 5.14-fold independent risk of cirrhosis in a nationally representative cohort, reinforcing the need for aetiology-specific diagnostic tools [22]. We acknowledge that the qualitative association between heavy drinking and cirrhosis is itself well established. The contribution of the present epidemiological component is therefore not the association per se, but (i) its quantification of a dose–response relationship in a recent, nationally representative U.S. cohort (NHANES 2017–2023), and (ii) its role as the population-level anchor that motivates, and is mechanistically linked—via CTD ethanol–gene evidence—to, the transcriptomic signature. The novelty of this work lies in the integration of population epidemiology, lncRNA–mRNA network analysis, multi-algorithm machine learning and computational drug repurposing within a single framework, rather than in any individual component. At the molecular level, transcriptomic profiling and multi-algorithm machine learning converged on a five-gene signature with an external validation AUC of 0.824, alongside two candidate therapeutics with favourable in silico binding profiles. It is important to note that the drug repurposing strategy employed here operates at the level of transcriptomic signature reversal rather than individual gene inhibition; accordingly, the molecular docking analyses of fluvastatin and honokiol against *IL1B* and *CCL3* reflect target engagement capacity and are interpreted in the context of the candidates' broader pleiotropic mechanisms, as discussed below.

A salient finding is the pronounced global downregulation of lncRNAs in ALC, with 91.7% exhibiting decreased expression. This contrasts sharply with the balanced distribution among mRNAs and suggests that widespread lncRNA suppression may constitute a candidate molecular feature of end-stage alcoholic liver disease requiring replication in larger independent cohorts. One plausible, though unproven, contributor is epigenetic silencing, as ethanol and acetaldehyde can perturb one-carbon metabolism and chromatin-modifying enzyme activity [23]; loss of functional hepatocyte mass in cirrhosis may further contribute, given the tissue-specific expression of many lncRNAs. We did not test these mechanisms directly and present them only as hypotheses to be examined in future epigenomic studies.

The Yellow and Pink modules provide mechanistic insight into the dual pathological processes driving ALC. The Yellow module was enriched for ECM organisation and integrin signalling, consistent with hepatic stellate cell activation in fibrogenesis [24]. Within this module, lncRNA *UCA1* emerged as a central hub connected to pro-fibrotic genes *SPP1* and *LAMA2*. *UCA1* has been implicated in HCC progression [25,26]. Within the limits of a co-expression analysis, our data position it as a hub correlated with pro-fibrotic genes, but the proposed microRNA-sponging mechanism is inferred from other contexts and was not assessed here. Conversely, the Pink module demonstrated enrichment for immune processes, including B cell receptor and NF- κ B signalling. The hub lncRNA *TMEM26-AS1* exhibited connectivity to inflammatory mediators *IL1B* and *CCL3*, as well as the ferroptosis pathway. The suppression of this module suggests immune dysfunction rather than simple inflammation, contributing to impaired hepatic immune surveillance in advanced disease [27]. Clinically, this distinction implies that immunostimulatory rather than immunosuppressive strategies may be more appropriate in the end-stage cirrhotic setting.

The five core diagnostic genes bridge the inflammatory and fibrotic arms of ALC pathogenesis. *IL1B* and *CCL3* are well-characterised inflammatory mediators [28,29]. More relevant to the present study, both transcripts are consistently down-regulated in ALC tissues across our training and validation cohorts—a seemingly paradoxical pattern that

we interpret below in the context of CAID. Their suppression in end-stage cirrhosis is biologically counterintuitive, yet aligns with emerging concepts of disease-stage-dependent immune remodelling. Our results specifically reflect the transcriptomic landscape of end-stage cirrhosis, a state increasingly recognised as Cirrhosis-Associated Immune Dysfunction (CAID), characterised by progressive immunodeficiency and systemic inflammation that may culminate in functional immune exhaustion [27]. This stands in contrast to early-stage alcohol-associated liver disease, particularly alcoholic hepatitis, which is characterised by a ‘cytokine storm’ and hyperinflammation [30]. Consequently, the suppression of *IL1B* and *CCL3* likely results from two converging mechanisms: the functional exhaustion of innate immune cells described in CAID [27], and the ‘burn-out’ effect caused by massive fibrotic replacement. Histologically, the accumulation of Type I collagen physically displaces functional hepatic parenchyma and resident immune niches [31,32], leading to a dilution of cytokine-producing cells within the dense fibrotic scar tissue. The fibrosis-related genes *LUM*, *SPP1* and *ITGA6* are established contributors to ECM remodelling and hepatic fibrosis [33,34]. Most relevant to our findings, *SPP1* has a reported stage-dependent, pro-fibrogenic role in advanced fibrosis, consistent with our observation that *SPP1* up-regulation maps to the fibrotic Yellow module. Our immune-infiltration analysis revealed an unexpected attenuation of macrophage-associated transcriptomic signatures rather than M2-dominant polarisation. We interpret this correlative observation cautiously: it is compatible with ECM-associated remodelling of the intrahepatic immune compartment, but a causal claim of physical immune-cell exclusion would require spatially resolved or single-cell data, which were not available here. We therefore present it as a hypothesis rather than an established mechanism.

The drug repurposing analysis nominated fluvastatin and honokiol as therapeutic candidates. Fluvastatin demonstrated strong binding affinity to *IL1B* (−7.65 kcal/mol). Beyond cholesterol-lowering effects, statins possess pleiotropic anti-inflammatory and anti-fibrotic properties through inhibition of RhoA and Rho-kinase signalling [35,36]. Recent clinical evidence demonstrates that statin use is associated with reduced cirrhosis progression and HCC risk in chronic liver disease patients [35,36]. Honokiol exhibited preferential binding to *CCL3* (−6.94 kcal/mol) and has demonstrated hepatoprotective effects. Recent experimental studies confirm that honokiol can inhibit epithelial-mesenchymal transition (EMT) and hepatic fibrosis by suppressing the Wnt/ β -catenin and TGF- β signalling pathways [37]. It warrants emphasis that the DSigDB-based repurposing approach identifies compounds capable of reversing the disease-associated transcriptomic signature as a whole, rather than exclusively targeting individual downregulated genes. Accordingly, the therapeutic relevance of fluvastatin and honokiol in the end-stage cirrhotic context is more plausibly attributable to their well-documented target-independent mechanisms—RhoA/Rho-kinase inhibition and TGF- β pathway suppression, respectively—than to direct modulation of *IL1B* or *CCL3*, which are suppressed at this disease stage. Prospective validation should therefore prioritise earlier fibro-inflammatory stages of ALC, where *IL1B* and *CCL3* are overexpressed and direct anti-inflammatory targeting remains mechanistically coherent. These compounds, identified through unbiased screening with established safety profiles, represent attractive candidates for stage-stratified evaluation. An important question is whether fluvastatin and honokiol modulate the expression or function of the signature genes. Mechanistically, the DSigDB query selects compounds predicted to reverse the disease-associated transcriptomic signature as a whole rather than to inhibit any single gene; the docking results for *IL1B* and *CCL3* (Figure 8) therefore index target-engagement capacity rather than the primary mode of action. To examine the question more directly, we performed a structured literature survey of reported drug–gene and drug–pathway interactions for both candidates against each signature gene. For the fibrotic arm, flu-

vastatin specifically attenuates hepatic stellate-cell activation and lowers α -SMA, NF- κ B and extracellular-matrix gene expression in the liver [38], consistent with the broader anti-fibrotic, RhoA/Rho-kinase-dependent actions of statins [35,36]; honokiol suppresses TGF- β 1/SMAD signalling and so attenuates the fibrotic ECM programme encompassing LUM, SPP1 and ITGA6 [39]. For the inflammatory arm, honokiol reduces hepatic IL-1 β through NF- κ B/NLRP3 inhibition in liver-injury models [40], whereas the relationship between statins and IL1B is context-dependent rather than uniformly inhibitory. We emphasise that no direct CCL3-expression datum is available for either compound; the CCL3 link rests on docking-based target engagement and broad NF- κ B/chemokine suppression, and is therefore presented as a prediction to be tested rather than an established effect. Accordingly, these remain computational and literature-based inferences, and direct experimental confirmation—for example, qPCR and Western-blot measurement of IL1B, CCL3, LUM, SPP1 and ITGA6 in drug-treated hepatic models—is a necessary next step, which we now explicitly flag as a limitation. Furthermore, the molecular docking results presented here represent static binding poses; molecular dynamics (MD) simulations would provide additional insight into the dynamic stability of these protein–ligand interactions and are identified as a priority for future computational validation.

Several limitations warrant consideration. Importantly, the present signature was derived and validated exclusively in liver-tissue transcriptomes and therefore constitutes a tissue-associated rather than a non-invasive biomarker set. Although several signature genes encode secreted or circulating proteins—LUM is an established serum fibrosis marker, and SPP1 (osteopontin), IL1B and CCL3 are detectable in peripheral blood—whether the tissue signature is recapitulated in serum or plasma remains untested. Translating these findings into a non-invasive assay will require paired profiling of tissue and circulating samples, ideally embedded within fibrosis-screening programmes that have been shown to support earlier diagnosis and alcohol abstinence [41]. In addition, the WGCNA co-expression modules and the initial feature-selection step were derived from a small training cohort ($n = 18$; 6 ALC and 12 controls), which may limit module stability and increase the risk of over-fitting. We sought to mitigate this in two ways: the final five-gene signature was retained only at the intersection of three independent algorithms (LASSO, Random Forest and SVM-RFE), and its discriminative performance was confirmed in a fully independent external cohort of 93 samples (AUC = 0.824) rather than being inferred from the training data alone. Nevertheless, formal module-preservation analysis in larger, multi-cohort datasets will be required to establish the robustness of the underlying co-expression structure, and the present modular assignments should be regarded as hypothesis-generating. A further consideration is whether the proposed signature would change with clinical status, particularly following alcohol abstinence. Our cross-sectional, tissue-based design cannot directly address this dynamic behaviour, and we acknowledge it as a limitation. Several lines of evidence nonetheless indicate that the signature has the biological potential to be status-responsive. First, our CTD analysis (Section 2.5) shows that four of the five genes (IL1B, CCL3, SPP1 and ITGA6) are directly ethanol-modulated, indicating that their expression tracks ethanol exposure. Second, sustained abstinence is the single most effective intervention in alcohol-associated liver disease and can drive partial regression of fibrosis, accompanied by reversion of disease-associated epigenetic and transcriptional states toward control [42,43]. Third, however, fibrosis resolution after cessation is non-uniform and can be actively restrained by persistent alcohol-induced epigenetic marks [43], implying that the fibrotic arm of the signature (LUM, SPP1, ITGA6) may decline only partially and variably after drinking stops. Integrating such molecular assessment with fibrosis screening and abstinence counselling, which has measurable clinical benefit [41], could provide a practical setting in which to test signature reversibility. Confirming whether the five-gene

signature is itself reversible will require longitudinal sampling of abstinent versus relapsing patients, an important direction for prospective validation. The present diagnostic signature was developed to distinguish ALC tissues from histologically normal liver tissues, rather than to differentiate ALC from cirrhosis of other aetiologies; future studies incorporating multi-aetiology cirrhosis cohorts will be necessary to evaluate the signature's aetiological specificity. The training cohort comprised a relatively small sample size ($n = 18$), and the retrospective nature precludes assessment of clinical variables such as disease duration and alcohol consumption patterns. While our analysis identified promising candidates, computational predictions require experimental validation in appropriate models before clinical translation. Future studies incorporating single-cell transcriptomics and prospective clinical validation will be essential to translate these findings into improved therapeutic approaches for ALC. Additionally, the immune infiltration analysis relied solely on ssGSEA; future work should employ complementary deconvolution methods such as CIBERSORT or xCell, and ideally single-cell or spatial transcriptomic data, to corroborate the observed macrophage signature attenuation. Finally, the drug repurposing candidates, fluvastatin and honokiol, were identified based on transcriptomic signature reversal in end-stage ALC tissues. Given that their primary diagnostic targets, IL1B and CCL3, are suppressed at this stage owing to CAID-associated immune exhaustion, their direct anti-inflammatory efficacy is most plausibly applicable to earlier, actively inflamed disease stages. This represents an inherent limitation of signature-based repurposing applied to end-stage specimens, and stage-specific experimental validation is warranted.

4. Materials and Methods

4.1. Study Population and Epidemiological Analysis

Data from 17,007 adults were extracted from the National Health and Nutrition Examination Survey (NHANES, 2017–2023) using the R package *nhanesA* [44]. Participants with viral hepatitis were excluded. Liver cirrhosis was defined by self-reported physician diagnosis, and alcohol consumption was stratified into never, non-heavy, and heavy drinking (≥ 4 –5 drinks/day). Survey-weighted logistic regression (adjusting for demographics, BMI, and diabetes) was employed to evaluate the epidemiological association between alcohol consumption and cirrhosis risk.

4.2. Transcriptomic Data Acquisition and Processing

The overall study design and analytical workflow, integrating epidemiological data with multi-cohort transcriptomics, are systematically illustrated in Figure 9. Gene expression data for ALC and normal liver tissues were obtained from the Gene Expression Omnibus (GEO) [45]. GSE142530 served as the training cohort ($n = 18$; ALC $n = 6$, normal controls $n = 12$). Three independent datasets (GSE103580, GSE28619, and GSE14323) were merged, rigorously corrected for batch effects using the ComBat function from the *sva* package (version 3.48.0) [46], and employed as the external validation cohort ($n = 93$). The combined validation cohort consisted of GSE103580 ($n = 40$), GSE28619 ($n = 22$), and GSE14323 ($n = 31$). Each dataset was individually normalised prior to merging using R software (version 4.5.1; R Foundation for Statistical Computing, Vienna, Austria) [47], with raw count data processed using the *data.table* package (version 1.17.8) [48]. Batch effect correction efficacy was confirmed via Principal Component Analysis (PCA) (Figure 10).

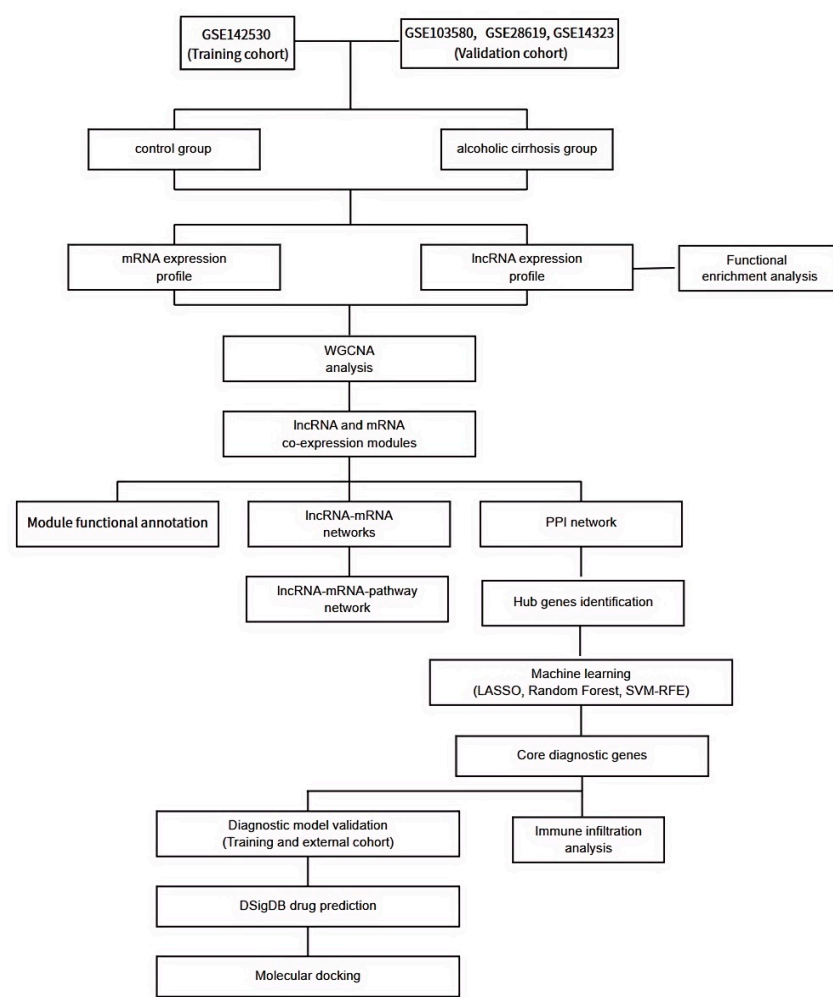


Figure 9. Overview of the study design. The overall study design and analytical workflow, integrating epidemiological data with multi-cohort transcriptomics.

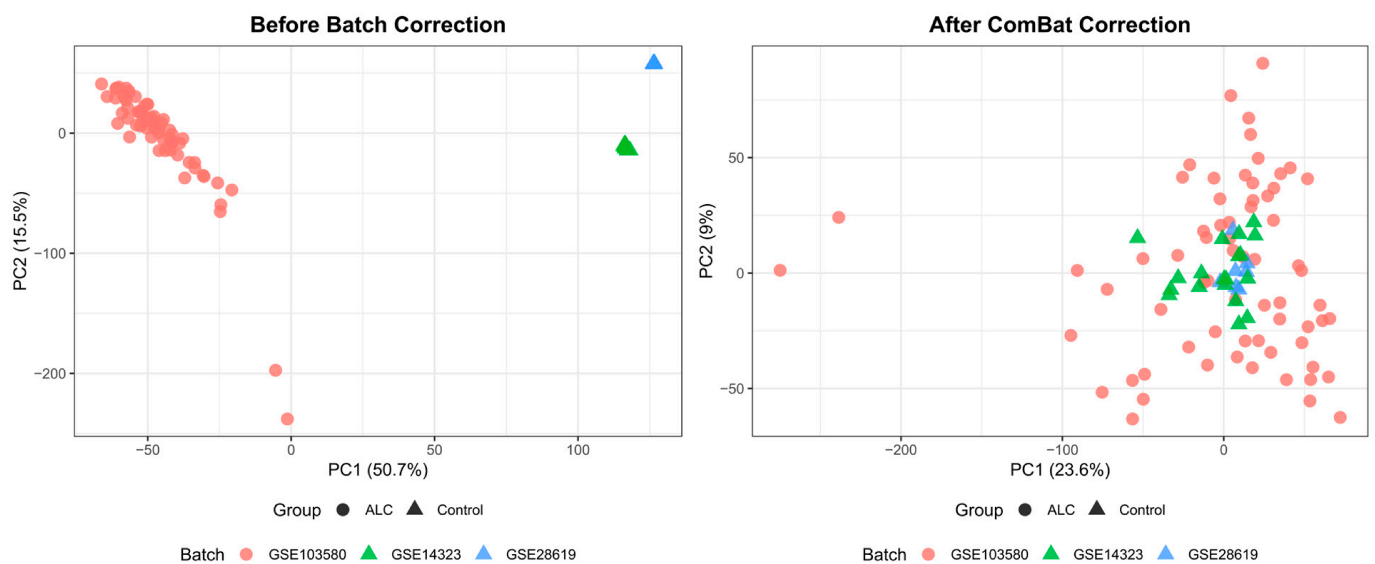


Figure 10. Principal Component Analysis (PCA) of the external validation cohort before and after batch effect correction. **(Left)** PCA plot before correction, showing strong batch-driven clustering. **(Right)** PCA plot after ComBat correction, demonstrating reduced batch effects and improved sample mixing.

4.3. Ethanol–Gene Interaction Validation via the Comparative Toxicogenomics Database

To establish a mechanistic bridge between the population-level epidemiological findings and the transcriptomic signature, the Comparative Toxicogenomics Database (CTD; ctdbase.org) was queried using the chemical term “Ethanol” against each of the five core diagnostic genes (*IL1B*, *CCL3*, *SPP1*, *ITGA6*, and *LUM*). The CTD curates literature-supported chemical–gene interactions, enabling systematic evaluation of whether the identified transcriptomic signature is biologically sensitive to ethanol exposure. Interaction evidence was classified as direct (ethanol shown to alter expression of the target gene or protein in published studies) or indirect (ethanol co-exposure with another chemical yielding the observed transcriptional alterations). This analysis serves to close the logical loop between the epidemiological observation of alcohol-driven cirrhosis risk and the molecular signature derived from ALC transcriptomes.

4.4. Co-Expression Network and Machine Learning Feature Selection

Differentially expressed genes (DEGs) were defined using thresholds of adjusted $p < 0.05$ and $|\log_2 \text{fold-change}| > 1$. Weighted Gene Co-expression Network Analysis (WGCNA) was performed using the WGCNA package [49] to isolate functional gene modules strongly correlated with ALC, with high-variability genes selected based on median absolute deviation (MAD), retaining the top 25% from the variance stabilising transformation (VST)-normalised expression matrix. Gene Ontology (GO) and KEGG pathway enrichment analyses were performed using the clusterProfiler package [50]. Protein–protein interaction (PPI) networks were constructed via the STRING database (minimum high-confidence interaction score threshold of 0.700) [51], and DEGs were imported into Cytoscape software (version 3.10.4) [52] for visualisation. Hub genes were identified using the CytoHubba plugin (Degree method) in Cytoscape [53]. Hub genes from the most significant modules were subjected to three parallel machine learning algorithms—LASSO regression using the glmnet package (version 4.1-7) [54], Random Forest (RF) implemented using the randomForest package (version 4.7-1.1) with 500 trees [55], and SVM-RFE conducted using the e1071 package (version 1.7-13) [56]—to extract a core diagnostic signature. Five-fold cross-validation was used to tune regularisation parameters in LASSO. Given the limited training sample size $n = 18$, the final combined diagnostic model was evaluated using leave-one-out cross-validation (LOOCV) to obtain a bias-corrected estimate of discriminative performance. The diagnostic endpoint was the binary classification of histologically confirmed ALC tissue versus normal liver tissue. Model performance was evaluated via the area under the receiver operating characteristic curve (AUC).

Detailed protocols, including data preprocessing and specific algorithmic configurations, are provided in Supplementary File S1 [57,58].

4.5. Immune Infiltration and Therapeutic Prediction

Immune cell abundance, with a specific focus on macrophage populations, was quantified using single-sample Gene Set Enrichment Analysis (ssGSEA) with immune cell gene sets derived from the MSigDB Hallmark and curated immune gene set collections (C7 collection). Immune cell abundance scores were computed per sample and compared between ALC and normal groups using the Wilcoxon rank-sum test. All statistical computations and visualisations were generated using the SciPy (version 1.10.1) [59], Seaborn (version 0.12.2) [60], and Matplotlib (version 3.7.1) [61] libraries in a Python 3.10 environment. Finally, the Drug Signatures Database (DSigDB) [62] was queried using the core gene signature to repurpose candidate therapeutics, which were subsequently validated through structural molecular docking simulations. Molecular docking was performed using AutoDock Vina (version 1.2.3; The Scripps Research Institute, La Jolla, CA, USA)

with a grid box centred on the active site of each target protein, for *IL1B* and *CCL3*, the two inflammatory hub genes with available high-resolution crystal structures in the Protein Data Bank; *ITGA6* was excluded from docking analysis due to the absence of a suitable ligand-bound crystal structure. Chemical structures of candidate compounds were obtained from the PubChem database [63]. Ligand geometries were subjected to energy minimisation via ChemBioOffice (PerkinElmer, Waltham, MA, USA) [64]. Receptor structures were retrieved from the RCSB Protein Data Bank [65]. Detailed protocols, including data preprocessing, software versions, and specific algorithmic configurations, are provided in Supplementary File S1.

5. Conclusions

This study provides integrated epidemiological and molecular evidence to advance the diagnosis and mechanistic understanding of alcoholic liver cirrhosis. Analysis of 17,007 adults from a nationally representative survey confirmed that heavy drinking independently confers a five-fold increased risk of cirrhosis in a dose-dependent manner. Transcriptomic profiling of human liver tissues revealed pronounced global suppression of long non-coding RNA expression, a finding that may reflect a candidate molecular feature of alcohol-related cirrhosis requiring replication in larger, multi-aetiology cohorts. A five-gene diagnostic signature (*IL1B*, *CCL3*, *LUM*, *SPP1*, *ITGA6*), derived from fibrosis- and immune-associated co-expression modules and developed to distinguish ALC from histologically normal liver tissue, achieved an area under the receiver operating characteristic curve of 0.824 in an independent external validation cohort of 93 samples. Immune microenvironment analysis uncovered global attenuation of macrophage-associated transcriptomic signatures rather than the expected M2-dominant polarisation, suggesting that immunological assessment may inform therapeutic strategy selection at different disease stages. Computational drug repurposing and molecular docking identified fluvastatin and honokiol as candidate therapeutic agents, with fluvastatin showing strong binding affinity to *IL1B* (−7.65 kcal/mol) and honokiol preferentially targeting *CCL3* (−6.94 kcal/mol). Collectively, these findings establish a molecular and epidemiological rationale for developing an ALC-specific diagnostic panel. Because the present signature is tissue-derived, its translation into a non-invasive (blood-based) assay represents an explicit goal for future validation. The analysis also nominates fluvastatin and honokiol as mechanistically plausible therapeutic leads, whose direct anti-inflammatory effects are most pertinent to earlier, actively inflamed stages of alcohol-associated liver disease, and whose anti-fibrotic pleiotropic actions may provide benefit across disease stages—both warranting prospective, stage-stratified clinical validation.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/ijms27135809/s1>.

Author Contributions: Conceptualisation, J.S. and H.W.; Methodology, H.W. and L.Z.; Software, H.W. and M.X.; Validation, H.W., W.D. and L.Z.; Formal Analysis, H.W. and M.X.; Investigation, H.W. and W.D.; Resources, J.S.; Data Curation, H.W. and M.X.; Writing—Original Draft Preparation, H.W.; Writing—Review & Editing, J.S. and W.D.; Visualisation, H.W. and M.X.; Supervision, J.S.; Project Administration, J.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Ethical review and approval were waived for this study. The epidemiological analysis utilised publicly available, de-identified data from the National Health and Nutrition Examination Survey (NHANES), which is exempt from institutional review board oversight. The transcriptomic analyses were conducted using publicly available datasets deposited

in the Gene Expression Omnibus (GEO; accession numbers GSE142530, GSE103580, GSE28619, and GSE14323), which do not involve direct human subjects research.

Informed Consent Statement: Not applicable. This study used only publicly available, de-identified secondary data from NHANES and GEO repositories; no direct patient recruitment or data collection was performed.

Data Availability Statement: The transcriptomic datasets analysed in this study are publicly available from the Gene Expression Omnibus (GEO) repository under the following accession numbers: GSE142530, GSE103580, GSE28619, and GSE14323. The epidemiological data are publicly available from the National Health and Nutrition Examination Survey (NHANES) at <https://www.cdc.gov/nchs/nhanes/> (accessed on 3 January 2026). All other data supporting the findings of this study are available within the article and its Supplementary Materials.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

ALC	Alcohol-associated liver cirrhosis
NHANES	National Health and Nutrition Examination Survey
GEO	Gene Expression Omnibus
DEG	Differentially expressed gene
lncRNA	Long non-coding RNA
mRNA	Messenger RNA
WGCNA	Weighted gene co-expression network analysis
LASSO	Least absolute shrinkage and selection operator
SVM-RFE	Support vector machine–recursive feature elimination
RF	Random Forest
AUC	Area under the receiver operating characteristic curve
LOOCV	Leave-one-out cross-validation
PPI	Protein–protein interaction
ssGSEA	Single-sample gene set enrichment analysis
ECM	Extracellular matrix
CTD	Comparative Toxicogenomics Database
DSigDB	Drug Signatures Database
MASLD	Metabolic dysfunction-associated steatotic liver disease
CAID	Cirrhosis-associated immune dysfunction
HCC	Hepatocellular carcinoma
PCA	Principal Component Analysis
BMI	Body mass index
OR	Odds ratio
CI	Confidence interval
FIB-4	Fibrosis-4 index
AST	Aspartate aminotransferase
ALT	Alanine aminotransferase
IQR	Interquartile range
TGF- β	Transforming growth factor-beta
EMT	Epithelial-mesenchymal transition
DNMT	DNA methyltransferase
HDAC	Histone deacetylase

References

1. Mackowiak, B.; Fu, Y.; Maccioni, L.; Gao, B. Alcohol-associated liver disease. *J. Clin. Investig.* **2024**, *134*, e176345. [[CrossRef](#)] [[PubMed](#)]
2. Wang, Y.; Wang, M.; Liu, C.; Hao, M.; Wang, W.; Li, Y.; Shi, J.; Jia, X.; Zhang, X.; Dang, S. Global burden of liver cirrhosis 1990–2019 and 20 years forecast: Results from the global burden of disease study 2019. *Ann. Med.* **2024**, *56*, 2328521. [[PubMed](#)]
3. Tan, Y.; Hu, Y.; Yang, Y.; Chu, H. Alcohol-Related Liver Disease and Metabolic Dysfunction-Associated Steatotic Liver Disease: Molecular Pathogenesis and Therapeutic Interventions. *MedComm (2020)* **2025**, *6*, e70532.
4. Singal, A.K.; Mathurin, P. Diagnosis and Treatment of Alcohol-Associated Liver Disease: A Review. *JAMA* **2021**, *326*, 165–176. [[CrossRef](#)] [[PubMed](#)]
5. Israelsen, M.; Rungratanawanich, W.; Thiele, M.; Liangpunsakul, S. Non-invasive tests for alcohol-associated liver disease. *Hepatology* **2024**, *80*, 1390–1407. [[PubMed](#)]
6. Zhang, L.; Wang, R.; Nan, Y.; Kong, L. Deciphering the role of LncRNA in alcoholic liver disease: Mechanisms and therapeutic potential. *Medicine* **2024**, *103*, e40378. [[CrossRef](#)] [[PubMed](#)]
7. Zhang, L.; Wang, R.; Nan, Y.; Kong, L. Molecular regulators of alcoholic liver disease: A comprehensive analysis of microRNAs and long non-coding RNAs. *Front. Med.* **2025**, *12*, 1482089.
8. Zhang, W.; Lu, W.; Jiao, Y.; Li, T.; Wang, H.; Wan, C. Identifying disease progression biomarkers in metabolic associated steatotic liver disease (MASLD) through weighted gene co-expression network analysis and machine learning. *J. Transl. Med.* **2025**, *23*, 472. [[CrossRef](#)] [[PubMed](#)]
9. Wu, X.; Zheng, X.; Ye, G. WGCNA combined with machine learning to explore potential biomarkers and treatment strategies for acute liver failure, with experimental validation. *liver* **2024**, *3*, 100133. [[CrossRef](#)] [[PubMed](#)]
10. Llamosas-Falcón, L.; Probst, C.; Buckley, C.; Jiang, H.; Lasserre, A.M.; Puka, K.; Tran, A.; Zhu, Y.; Rehm, J. How does alcohol use impact morbidity and mortality of liver cirrhosis? A systematic review and dose-response meta-analysis. *Hepatol. Int.* **2024**, *18*, 216–224. [[PubMed](#)]
11. Yao, H.; Zhang, D.; Yu, H.; Shen, H.; Lan, X.; Liu, H.; Chen, X.; Wu, X.; Zhang, G.; Wang, X. AMPAkinase CX516 alleviated chronic ethanol exposure-induced neurodegeneration and depressive-like behavior in mice. *Toxicol. Appl. Pharmacol.* **2022**, *439*, 115924. [[CrossRef](#)] [[PubMed](#)]
12. Tiwari, V.; Chopra, K. Resveratrol abrogates alcohol-induced cognitive deficits by attenuating oxidative-nitrosative stress and inflammatory cascade in the adult rat brain. *Neurochem. Int.* **2013**, *62*, 861–869. [[PubMed](#)]
13. Singh, A.K.; Jiang, Y.; Benlhabib, E.; Gupta, S. Herbal mixtures consisting of puerarin and either polyenylphosphatidylcholine or curcumin provide comprehensive protection against alcohol-related disorders in P rats receiving free choice water and 15% ethanol in pure water. *J. Med. Food* **2007**, *10*, 526–542. [[CrossRef](#)] [[PubMed](#)]
14. Mukhopadhyay, P.; Horváth, B.; Rajesh, M.; Varga, Z.V.; Gariani, K.; Ryu, D.; Cao, Z.; Holovac, E.; Park, O.; Zhou, Z.; et al. PARP inhibition protects against alcoholic and non-alcoholic steatohepatitis. *J. Hepatol.* **2017**, *66*, 589–600. [[CrossRef](#)] [[PubMed](#)]
15. Magierowska, K.; Magierowski, M.; Hubalewska-Mazgaj, M.; Adamski, J.; Surmiak, M.; Sliwowski, Z.; Kwiecién, S.; Brzozowski, T. Carbon Monoxide (CO) Released from Tricarbonyldichlororuthenium (II) Dimer (CORM-2) in Gastroprotection against Experimental Ethanol-Induced Gastric Damage. *PLoS ONE* **2015**, *10*, e0140493. [[CrossRef](#)] [[PubMed](#)]
16. Glover, M.; Cheng, B.; Fan, R.; Pruett, S. The role of stress mediators in modulation of cytokine production by ethanol. *Toxicol. Appl. Pharmacol.* **2009**, *239*, 98–105. [[CrossRef](#)] [[PubMed](#)]
17. Gaballah, H.H.; Ghanem, H.B.; Tahoon, N.M.; Mohamed, D.A.; Ebeid, A.M. Hesperidin promotes lysosomal biogenesis in chronically ethanol-induced cardiotoxicity in rats: A proposed mechanisms of protection. *J. Biochem. Mol. Toxicol.* **2019**, *33*, e22253. [[PubMed](#)]
18. Lee, J.H.; Banerjee, A.; Ueno, Y.; Ramaiah, S.K. Potential relationship between hepatobiliary osteopontin and peroxisome proliferator-activated receptor alpha expression following ethanol-associated hepatic injury in vivo and in vitro. *Toxicol. Sci.* **2008**, *106*, 290–299. [[PubMed](#)]
19. Abdel Hamid, O.I.; Sabik, L.M.E.; Abdelfadeel, K.F.; Shaban, S.F. Tramadol aggravates cardiovascular toxicity in a rat model of alcoholism: Involvement of intermediate microfilament proteins and immune-expressed osteopontin. *J. Biochem. Mol. Toxicol.* **2021**, *35*, 1–15. [[CrossRef](#)] [[PubMed](#)]
20. O'Brien, M.A.; Weston, R.M.; Sheth, N.U.; Bradley, S.; Bigbee, J.; Pandey, A.; Williams, R.W.; Wolstenholme, J.T.; Miles, M.F. Ethanol-Induced Behavioral Sensitization Alters the Synaptic Transcriptome and Exon Utilization in DBA/2J Mice. *Front. Genet.* **2018**, *9*, 402. [[CrossRef](#)] [[PubMed](#)]
21. Furuya, S.; A Cichocki, J.; Konganti, K.; Dreval, K.; Uehara, T.; Katou, Y.; Fukushima, H.; Kono, H.; Pogribny, I.P.; Argemi, J.; et al. Histopathological and Molecular Signatures of a Mouse Model of Acute-on-Chronic Alcoholic Liver Injury Demonstrate Concordance With Human Alcoholic Hepatitis. *Toxicol. Sci.* **2019**, *170*, 427–437. [[PubMed](#)]
22. Jophlin, L.L.; Singal, A.K.; Bataller, R.; Wong, R.J.; Sauer, B.G.; Terrault, N.A.; Shah, V.H. ACG Clinical Guideline: Alcohol-Associated Liver Disease. *Am. J. Gastroenterol.* **2024**, *119*, 30–54. [[PubMed](#)]

23. Ajoolabady, A.; Aslkhodapasandhokmabad, H.; Zhou, Y.; Ren, J. Epigenetic modification in alcohol-related liver diseases. *Med. Res. Rev.* **2022**, *42*, 1463–1491. [[CrossRef](#)] [[PubMed](#)]
24. Patidar, P.; Hirani, N.; Bharti, S.; Baig, M.S. Key regulators of hepatic stellate cell activation in alcohol liver Disease: A comprehensive review. *Int. Immunopharmacol.* **2024**, *141*, 112938. [[CrossRef](#)] [[PubMed](#)]
25. Liang, W.; Zhao, Y.; Meng, Q.; Jiang, W.; Deng, S.; Xue, J. The role of long non-coding RNA in hepatocellular carcinoma. *Aging* **2024**, *16*, 4052–4073. [[CrossRef](#)] [[PubMed](#)]
26. Yang, Z.; Zhang, H.; Yin, M.; Cheng, Z.; Jiang, P.; Feng, M.; Liu, Z.; Liao, B. TGF- β 1/Smad3 upregulates UCA1 to promote liver fibrosis through DKK1 and miR18a. *J. Mol. Med.* **2022**, *100*, 1465–1478. [[CrossRef](#)] [[PubMed](#)]
27. Albillos, A.; Martin-Mateos, R.; Van der Merwe, S.; Wiest, R.; Jalan, R.; Álvarez-Mon, M. Cirrhosis-associated immune dysfunction. *Nat. Rev. Gastroenterol. Hepatol.* **2022**, *19*, 112–134. [[PubMed](#)]
28. Hammerich, L.; Tacke, F. Hepatic inflammatory responses in liver fibrosis. *Nat. Rev. Gastroenterol. Hepatol.* **2023**, *20*, 633–646. [[CrossRef](#)] [[PubMed](#)]
29. Louvet, A.; Mathurin, P. Alcoholic liver disease: Mechanisms of injury and targeted treatment. *Nat. Rev. Gastroenterol. Hepatol.* **2015**, *12*, 231–242. [[CrossRef](#)] [[PubMed](#)]
30. Gao, B.; Bataller, R. Alcoholic liver disease: Pathogenesis and new therapeutic targets. *Gastroenterology* **2011**, *141*, 1572–1585. [[CrossRef](#)] [[PubMed](#)]
31. Acharya, P.; Chouhan, K.; Weiskirchen, S.; Weiskirchen, R. Cellular Mechanisms of Liver Fibrosis. *Front. Pharmacol.* **2021**, *12*, 671640. [[CrossRef](#)] [[PubMed](#)]
32. Kisseleva, T.; Brenner, D. Molecular and cellular mechanisms of liver fibrosis and its regression. *Nat. Rev. Gastroenterol. Hepatol.* **2021**, *18*, 151–166. [[PubMed](#)]
33. Li, B.; Hu, J.; Xu, H. Integrated single-cell and bulk RNA sequencing reveals immune-related SPP1+ macrophages as a potential strategy for predicting the prognosis and treatment of liver fibrosis and hepatocellular carcinoma. *Front. Immunol.* **2024**, *15*, 1455383. [[CrossRef](#)] [[PubMed](#)]
34. Jin, H.; Kim, W.; Yuan, M.; Li, X.; Yang, H.; Li, M.; Shi, M.; Turkez, H.; Uhlen, M.; Zhang, C.; et al. Identification of SPP1 (+) macrophages as an immune suppressor in hepatocellular carcinoma using single-cell and bulk transcriptomics. *Front. Immunol.* **2024**, *15*, 1446453. [[CrossRef](#)] [[PubMed](#)]
35. Sharpton, S.R.; Loomba, R. Emerging role of statin therapy in the prevention and management of cirrhosis, portal hypertension, and HCC. *Hepatology* **2023**, *78*, 1896–1906. [[CrossRef](#)] [[PubMed](#)]
36. Kraglund, F.; Christensen, D.H.; Eiset, A.H.; Villadsen, G.E.; West, J.; Jepsen, P. Effects of statins and aspirin on HCC risk in alcohol-related cirrhosis: Nationwide emulated trials. *Hepatol. Commun.* **2023**, *7*, e0013. [[CrossRef](#)] [[PubMed](#)]
37. Yang, C.; Zhao, Y.; Luo, Z.; Hu, Y.; Wang, S.; Hu, S.; Yao, Y.; Pan, L.; Shen, C.P.; Xu, T. Honokiol Inhibits the Inflammatory Response and Lipid Metabolism Disorder by Inhibiting p38 α in Alcoholic Liver Disease. *Planta Medica* **2023**, *89*, 273–285. [[PubMed](#)]
38. Chong, L.W.; Hsu, Y.C.; Lee, T.F.; Lin, Y.; Chiu, Y.-T.; Yang, K.-C.; Wu, J.-C.; Huang, Y.-T. Fluvastatin attenuates hepatic steatosis-induced fibrogenesis in rats through inhibiting paracrine effect of hepatocyte on hepatic stellate cells. *BMC Gastroenterol.* **2015**, *15*, 22. [[CrossRef](#)] [[PubMed](#)]
39. Kataoka, S.; Umemura, A.; Okuda, K.; Taketani, H.; Seko, Y.; Nishikawa, T.; Yamaguchi, K.; Moriguchi, M.; Kanbara, Y.; Arbisser, J.L.; et al. Honokiol Acts as a Potent Anti-Fibrotic Agent in the Liver through Inhibition of TGF- β 1/SMAD Signaling and Autophagy in Hepatic Stellate Cells. *Int. J. Mol. Sci.* **2021**, *22*, 13354. [[CrossRef](#)] [[PubMed](#)]
40. Alhussaini, A.F.; Elshal, M.; Hazem, S.H. Honokiol Ameliorates LPS/D-GalN-Induced Acute Liver Failure via Activation of SIRT3/AMPK and Keap1/Nrf2/HO-1 Signaling and Inhibition of the NF- κ B/NLRP3 Inflammasome Axis. *Pharmaceuticals* **2026**, *19*, 909.
41. Avitabile, E.; Gratacós-Ginès, J.; Pérez-Guasch, M.; Rubio, A.B.; Herms, Q.; Cervera, M.; Nadal, R.; Carol, M.; Fabrellas, N.; Bruguera, P.; et al. Liver fibrosis screening increases alcohol abstinence. *JHEP Rep.* **2024**, *6*, 101165. [[CrossRef](#)] [[PubMed](#)]
42. Diaz, L.A.; Winder, G.S.; Leggio, L.; Bajaj, J.S.; Bataller, R.; Arab, J.P. New insights into the molecular basis of alcohol abstinence and relapse in alcohol-associated liver disease. *Hepatology* **2025**, *82*, 254–271. [[CrossRef](#)] [[PubMed](#)]
43. Schonfeld, M.; O’neil, M.; Weinman, S.A.; Tikhanovich, I. Alcohol-induced epigenetic changes prevent fibrosis resolution after alcohol cessation in miceresolution. *Hepatology* **2024**, *80*, 119–135. [[PubMed](#)]
44. Choday, S.; Zahdeh, T.; Kang, P.; Reynolds, J.; Wong, R. Insights on Alcohol-Associated Liver Disease, a Decade of Data from National Survey. *Gastrointest. Disord.* **2025**, *7*, 52. [[CrossRef](#)]
45. Clough, E.; Barrett, T.; Wilhite, S.E.; Ledoux, P.; Evangelista, C.; Kim, I.F.; Tomashevsky, M.; A Marshall, K.; Phillippy, K.H.; Sherman, P.M.; et al. NCBI GEO: Archive for gene expression and epigenomics data sets: 23-year update. *Nucleic Acids Res.* **2024**, *52*, D138–D144. [[PubMed](#)]
46. Leek, J.T.; Johnson, W.E.; Parker, H.S.; Jaffe, A.E.; Storey, J.D. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* **2012**, *28*, 882–883. [[CrossRef](#)] [[PubMed](#)]

47. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2025.
48. Dowle, M.; Srinivasan, A. data.table: Extension of data.frame, R Package Version 1.17.8. 2024. Available online: <https://www.r-project.org/> (accessed on 22 June 2026).
49. Langfelder, P.; Horvath, S. WGCNA: An R package for weighted correlation network analysis. *BMC Bioinform.* **2008**, *9*, 559. [[CrossRef](#)]
50. Wu, T.; Hu, E.; Xu, S.; Chen, M.; Guo, P.; Dai, Z.; Feng, T.; Zhou, L.; Tang, W.; Zhan, L.; et al. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innovation* **2021**, *2*, 100141. [[CrossRef](#)] [[PubMed](#)]
51. Szklarczyk, D.; Nastou, K.; Koutrouli, M.; Kirsch, R.; Mehryary, F.; Hachilif, R.; Hu, D.; Peluso, M.E.; Huang, Q.; Fang, T.; et al. The STRING database in 2025: Protein networks with directionality of regulation. *Nucleic Acids Res.* **2025**, *53*, D730–D737. [[PubMed](#)]
52. Shannon, P.; Markiel, A.; Ozier, O.; Beliga, N.S.; Wang, J.T.; Ramage, D.; Amin, N.; Schwikowski, B.; Ideker, T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res.* **2003**, *13*, 2498–2504. [[CrossRef](#)] [[PubMed](#)]
53. Chin, C.H.; Chen, S.H.; Wu, H.H.; Ho, C.-W.; Ko, M.-T.; Lin, C.-Y. cytoHubba: Identifying hub objects and sub-networks from complex interactome. *BMC Syst. Biol.* **2014**, *8*, S11. [[PubMed](#)]
54. Tay, J.K.; Narasimhan, B.; Hastie, T. Elastic net regularization paths for all generalized linear models. *J. Stat. Softw.* **2023**, *106*, 1–31. [[CrossRef](#)] [[PubMed](#)]
55. Breiman, L. Random forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
56. Meyer, D.; Dimitriadou, E.; Hornik, K.; Weingessel, A.; Leisch, F. e1071: Misc Functions of the Department of Statistics, Probability Theory Group (Formerly: E1071), TU Wien, R Package Version 1.7-13. 2019. Available online: <https://www.r-project.org/> (accessed on 22 June 2026).
57. Ma, C.; Zhang, X.; Zhang, W.; Duan, J.; Yang, H. Association between serum homocysteine levels and advanced hepatic fibrosis in alcohol-related liver disease: A cross-sectional study of NHANES. *Medicine* **2025**, *104*, e43395. [[PubMed](#)]
58. Li, X.M.; Liu, S.L.; He, Y.J.; Shu, J.C. Using new indices to predict metabolism dysfunction-associated fatty liver disease (MAFLD): Analysis of the National Health and Nutrition Examination Survey database. *BMC Gastroenterol.* **2024**, *24*, 109. [[CrossRef](#)] [[PubMed](#)]
59. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J. SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nat. Methods* **2020**, *17*, 261–272. Erratum in *Python. Nat. Methods* **2020**, *17*, 352. [[CrossRef](#)] [[PubMed](#)]
60. Waskom, M.L. seaborn: Statistical data visualization. *J. Open Source Softw.* **2021**, *6*, 3021. [[CrossRef](#)]
61. Hunter, J.D. Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [[CrossRef](#)]
62. Yoo, M.; Shin, J.; Kim, J.; Kim, J.; Ryall, K.A.; Lee, K.; Lee, S.; Jeon, M.; Kang, J.; Tan, A.C. DSigDB: Drug signatures database for gene set analysis. *Bioinformatics* **2015**, *31*, 3069–3071. [[CrossRef](#)] [[PubMed](#)]
63. Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Scoemaker, B.A.; Thiessen, P.A.; Yu, B.; et al. PubChem 2025 update. *Nucleic Acids Res.* **2025**, *53*, D1516–D1525. [[PubMed](#)]
64. Kerwin, S.M. ChemBioOffice Ultra 2010 suite. *J. Am. Chem. Soc.* **2010**, *132*, 2466–2467. [[PubMed](#)]
65. Burley, S.K.; Bhikadiya, C.; Bi, C.; Bittrich, S.; Chao, H.; Craig, P.A.; Crichlow, G.V.; Dalenberg, K.; Duarte, J.M.; Dutta, S.; et al. RCSB Protein Data Bank (RCSB.org): Delivery of experimentally-determined PDB structures alongside one million computed structure models of proteins from artificial intelligence/machine learning. *Nucleic Acids Res.* **2023**, *51*, D488–D508. [[PubMed](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.