



Analysis

Unveiling the role of IL7R in metabolism-associated fatty liver disease leading to hepatocellular carcinoma through transcriptomic and machine learning approaches

Priyadharshini Annadurai¹ · Arnold Emerson Isaac¹ 

Received: 23 January 2025 / Accepted: 9 May 2025

Published online: 23 May 2025

© The Author(s) 2025 

Abstract

Dysregulation of hepatic metabolism is a crucial factor in the development of fatty liver disease and significantly increases the risk of hepatocellular carcinoma (HCC). This study aims to identify the genes implicated in the prognosis of HCC among individuals suffering from metabolic fatty liver disease. We analysed protein–protein interaction (PPI) networks and constructed a weighted gene co-expression network analysis (WGCNA) using high-throughput gene expression profiling datasets. Our meta-analysis uncovered 442 differentially expressed genes (DEGs), comprising 30 upregulated and 412 downregulated genes. We constructed a PPI network from the DEGs and identified significant hub genes based on their degree centrality scores. Additionally, WGCNA highlighted impactful genes and tightly correlated modules, leading to the creation of a gene interaction network specific to metabolism-associated fatty liver disease (MAFLD). Pathway analysis revealed the candidate regulatory gene interleukin-7 receptor (IL7R), which is involved in cytokine-mediated signalling across both interaction networks. Pro-inflammatory cytokines interact with IL7R, activating the JAK/STAT pathway that influences gene expression throughout progression to HCC. IL7R activates STAT3, affecting the behaviour of activated hepatic stellate cells following initial liver damage. Furthermore, the expression of the IL7R gene was validated as a predictor of HCC malignancy through a logistic regression model, resulting in an accuracy of 92%. Findings suggest that IL7R could be the target gene associated with metabolism-linked HCC. It could significantly impact the management of metabolic-associated fatty liver disease (MAFLD) and may help enhance HCC diagnostics to improve patient outcomes.

Keywords Metabolic-associated fatty liver disease · Hepatocellular carcinoma · Steatosis · Protein–protein interaction · Hepatic stellate cells · Cytokine signaling

1 Introduction

Metabolic-associated fatty liver disease (MAFLD) prevalence has risen worldwide, making it the most common liver disease, affecting more than one-third of adults [1]. In 2020, a new term, MAFLD, was proposed for Non-alcoholic fatty liver disease (NAFLD), based on hepatic steatosis and at least one of type 2 diabetes mellitus (T2D), obesity, or metabolic dysregulation [2–4]. NAFLD is a chronic liver disease with no obvious cause, making it an exclusion diagnosis where other

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s12672-025-02638-5>.

✉ Arnold Emerson Isaac, i_arnoldemerson@yahoo.com | ¹Bioinformatics Programming Laboratory, Department of Bioscience, School of Bio Science and Technology, Vellore Institute of Technology, Katpadi, Vellore - 632014, Tamil Nadu, India.



possible conditions were ruled out [5]. NAFLD affects 38% of the global population, with a growing risk of severe complications [6]. The disease is becoming more prevalent at an earlier age due to changes in diet, urbanization, obesity, and T2D. Also, there are some geographical disparities in NAFLD prevalence rates partly due to socio-economic and genetic factors [7]. Research has shown that NAFLD patients are more prone to develop hepatocellular carcinoma (HCC) [8, 9]. The burden of HCC is the third leading cause of cancer-related death, and one of the major causes is due to MAFLD [10–12]. Recently, MAFLD-related HCC has been reported more often, while viral hepatitis-related HCC has declined over the years [13]. Also, it disrupts normal mechanisms of hepatic stellate cells (HSCs), which are crucial in liver injury repair [14]. NAFLD is becoming increasingly recognized due to its rising prevalence worldwide, greater awareness among healthcare professionals, and improved understanding of these disorders [15, 16]. Patients with NAFLD and alcoholic liver disease are at a higher risk of developing HCC compared to those with viral infections such as hepatitis B and hepatitis C [17].

The presence of lipid droplets or lipid bodies in the organelles of hepatocytes is a characteristic feature of NAFLD, which can progress to fibrosis and HCC [18]. Hepatic steatosis, which is characterized by the accumulation of fat in the liver, is the hallmark of NAFLD etiology, and it is an evolving condition that can progress to cirrhosis and liver cancer [19, 20]. Even patients with treated and maintained virological responses are at a higher risk of developing HCC due to steatosis. The risk could be up to six times higher in such cases [21]. Obesity, hyperinsulinemia, glycaemic index (GI), and diabetes are some of the common risk factors for both steatosis and liver cancer. These factors can alter liver microcirculation and inflammatory cytokines, leading to liver metastases [22, 23]. The procoagulant state in MAFLD patients increases with NAFLD severity, possibly due to liver fat content, inflammation, atherogenic dyslipidemia, obesity or insulin resistance [24]. Likewise, the intake of imbalanced macronutrients such as saturated fatty acids (SFA), trans fats, and simple sugars can contribute to the development of NAFLD, which can have detrimental effects on liver metabolism [25, 26]. Moreover, hepatic steatosis is also linked to increased post-operative morbidity and mortality in individuals undergoing liver surgery [27]. Additionally, alcohol consumption can lead to diverse types of alcoholic liver disease (ALD), including hepatic steatosis, alcoholic steatohepatitis (ASH), fibrosis, cirrhosis, and HCC [28, 29]. Although some studies have suggested a potential association between hepatitis C virus (HCV) and alcoholism with HCC development, the causal relationship between HCV and alcoholic HCC remains controversial [30].

The pathogenesis of NAFLD-related liver carcinogenesis is distinct and involves a complex interplay of multiple processes, including environmental factors, oxidative stress, chronic inflammation, and the immunological response. A meta-analysis reveals that NAFLD, characterized by abnormal liver enzyme levels and abdominal ultrasound, increases the risk of T2D and metabolic syndrome, necessitating non-invasive procedures for treatment [31]. MAFLD, when combined with other liver diseases, increases complications and mortality rates in patients, while NAFLD research focusing on hepatic phenotype is much needed [32]. Based on the literature survey, the present study identified the predictive biomarkers involved in developing HCC with steatosis. This study considered the development of HCC with fatty liver disease, metabolic disorders, alcohol abuse, and steatosis as underlying disorders that fulfil the MAFLD criteria [33, 34]. Transcriptomic data of the patients were used, and differential gene expression analysis was carried out after normalizing the samples. Functional enrichment and pathway analysis were performed to explore the function of DEGs. The protein–protein interaction (PPI) network was constructed to identify the hub genes. Additionally, weighted gene co-expression network analysis (WGCNA) was carried out with a validation dataset to find the module-trait relationship of the prominent biomarker genes. WGCNA helps researchers identify functional interpretation of the genes that is biologically meaningful in uncovering cancer progression processes, and correlate gene modules with clinical data for diagnostic and prognostic biomarkers, highlighting key regulatory genes as therapeutic targets [35–37]. In the WGCNA module, the co-expression network was built, and the correlation between each pair of genes was computed initially. Then, their distribution was fitted to a power law, and an adjacency score was calculated accounting for the network topology. Based on these approaches prominent biomarkers were identified for the prognosis of HCC underlying MAFLD.

2 Methods

2.1 Dataset acquisition

The study utilized two RNA-seq datasets that were obtained from the publicly available gene expression omnibus (GEO) database [38]. The search strategy was as follows: “fatty liver disease”, “homo sapiens” (organism), and “expression profiling by high throughput sequencing” (filter). The resulting entries were further screened based on the relevance to HCC, not exposed to viral infection and non-treatment-based studies such as drug treatments. The first

dataset GSE140462 consisted of 14 samples, out of which 7 samples had a fatty liver background, and the remaining 7 were adjacent normal tissues of the patients [39]. The second dataset GSE184733 comprised 34 samples, of which 6 samples were considered for the study, containing 3 in each patient and normal samples, while the other samples were excluded due to viral infections [40].

2.2 Gene expression fold-change analysis

To ensure accurate results, the study filtered out lowly expressed genes as they could increase false discovery rates (FDR) and limit the ability to detect differentially expressed genes (DEGs) while estimating the FDR, where the rate of those features called significant is truly null. The edgeR function was used to filter these genes based on minimum counts per million (CPM) thresholds [41]. Samples with CPM above 0.5 in at least two sample groups were retained. Normalization was performed to eliminate composition biases between samples and the quality control of the samples was further visualized using MDS plot and boxplot (Fig. S1). To prepare the data for differential gene expression analysis, the read counts were transformed into logCPM, taking into account the mean–variance relationship in the data [42]. After converting the voom transformation, the study used the limma function to analyze the DEGs. The genes were filtered based on the \log_2 fold change (FC) cutoff of 1 and an adjusted p-value (Benjamin-Hochberg) less than 0.05.

2.3 Network analysis

The PPI network of the differentially expressed genes was constructed using the STRING database (<https://string-db.org/>) [43] with a confidence level of 0.4 and higher as a filter. The network was visualized using Cytoscape [44, 45]. The top 10 genes were identified based on their degree as hub genes using the Cytohubba plugin, as described by Bader & Hogue (2003) and Chin et al. (2014) [46, 47]. The degree centrality was calculated based on the number of neighbors a node has, and it determines the dominance of the network graph using Freeman's general formula, indicating the degree of centralization [48, 49].

$$C_D = \frac{\sum_{i=1}^n (k_{max} - k_i)}{(n-1)(n-2)}$$

where C_D is the degree centralization of the network, n determines the number of nodes, k_i defines the node i , and k_{max} is the maximum degree in the network, this formula calculates the sum of the difference between the maximum degree and the degree of each node, then normalize it by dividing by the maximum possible sum of differences. The sum considers degrees between the largest and given vertex, while the denominator normalizes results. The degree determines the number of interactions associated with the particular node, and hub genes have a large degree of interactions, and they are hypothesized to play an essential role in the network, which is associated with the biological process [50, 51].

2.4 Co-expression network analysis

The study employed GSE115193, an independent validation dataset, to construct a gene co-expression network using the WGCNA (v 1.47) package in R [52]. After filtering lowly expressed genes and normalization, gene expression values were imported into WGCNA to construct co-expression modules using the automatic network construction function `blockwiseModules` with default settings, assigning the soft threshold power and Topological Overlap Matrix (TOM) Type was unsigned. Then, the Pearson correlation analysis was used to identify modules significantly associated with phenotypic characteristics [53]. Finally, the modules network was mapped using Cytoscape 3.7.2. The overall pipeline of the study has been shown in Fig. 1 to find the driver genes.

2.5 Pathway enrichment analysis

Functional and enrichment analysis of the hub genes involved in the biological process, cellular components, molecular functions and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis was carried out using the DAVID (Database for Annotation, Visualization and Integrated Discovery) online tool [54]. The Benjamin-Hochberg method was used to calculate the p-value [55].

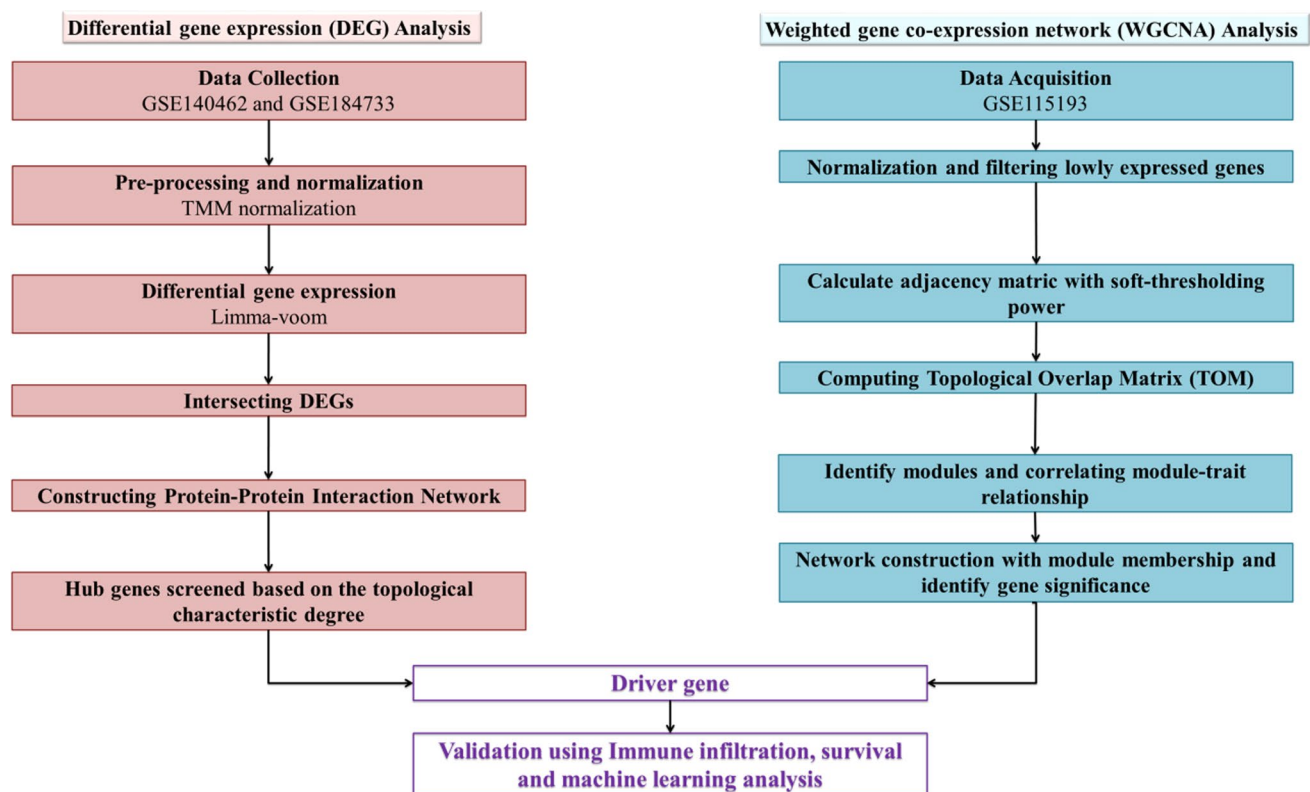


Fig. 1 Study workflow of this integrated bioinformatics analysis

2.6 Validation of hub gene

The hub genes were further studied for immune infiltration and survival analysis. Also, logistic regression machine learning models were used to predict the HCC outcome using the hub gene expression pattern through the TCGA-LIHC dataset [56]. The logistic regression model is preferred for predicting binary outcomes, particularly gene expression levels, due to its ease of modelling, robustness, and ability to include categorical variables, even in small datasets [57]. The model was validated based on the 10-fold cross-validation, model performance, precision and recall.

3 Results

3.1 Differential gene expression analysis

After normalization, the variance between the sample groups has been studied. The Euclidean distances were calculated, and the distance between each pair of points represents the variance between the sample groups [58, 59]. The MDS plot shows the variance between the samples based on their dimensions, with the x-axis being Dimension 1 and the y-axis being Dimension 2. In Fig. 2A, dimension 1 is 21%, and dimension 2 is 16% in the samples from the GSE140462. Similarly, in Fig. 2B, dimension 1 of GSE184733 is 38%, and dimension 2 is 25%, which shows the variance among the sample group is lesser when compared to the first dimension. In this analysis, both datasets showed higher first-dimensional variance when compared to dimension 2.

After transformation, the genes are analyzed for differential expression using the Limma tool [60]. The empirical Bayes function was used to predict the up and down-regulated genes, with a fold change cutoff of 1 and a p-value less than 0.05 (Fig. 2C, D). This approach helps in identifying the genes that are significantly differentially expressed between the two datasets, which can be further studied to understand their role in the biological processes of interest. A Venn diagram approach was used to identify the common DEGs in both datasets [61]. The analysis revealed that 30 commonly

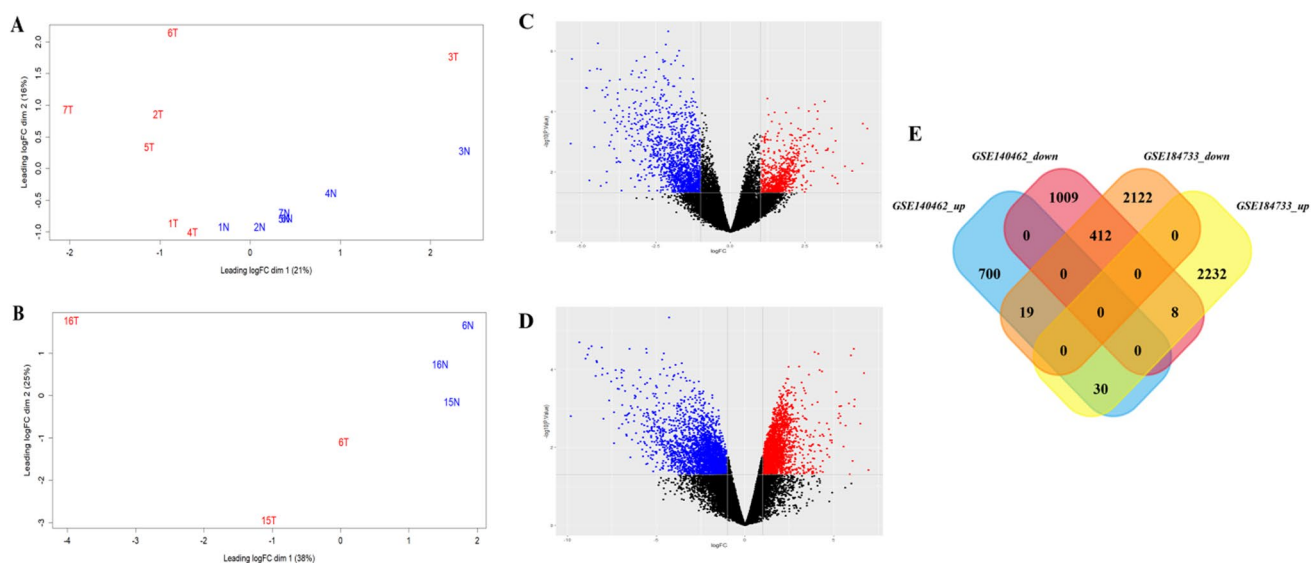


Fig. 2 RNA-Seq differential gene expression. **A** MDS plot for GSE140462. **B** MDS plot for GSE184733 samples. T represents the cancer samples, and N represents the normal samples. **C** Volcano plot for GSE140462. **D** Volcano plot for GSE184733. **E** Venn diagram identify the intersecting DEGs in both datasets

up-regulated and 412 commonly down-regulated genes were expressed in both datasets (Fig. 2E). These genes were further subjected to protein–protein interaction (PPI) network analysis.

3.2 Network analysis

A PPI network was constructed using 442 genes from the DEG analysis with the help of the STRING database. The resulting PPI network consisted of 338 nodes and 756 edges (Fig. 3A). The Cytohubba plugin was used to identify the top 10 hub genes based on their degree centrality as their topological characteristics (Table 1).

Based on Freeman's algorithm, the hub genes were screened and highlighted in Fig. 3A. The hub genes resulted from the PPI network analysis include IL1B, CXCL12, PLEK, HGF, NCAM1, FOS, FCGR3B, LCP2, CD274 and IL7R. The enriched terms of the hub genes are shown in Fig. 3B, which are involved in the biological pathways. Also, the detailed description of the enrichment analysis has been represented in the Table.S2.

3.3 Weighted gene correlation network analysis

The WGCNA was used to analyze the connection between genes and physiological traits, discovering the hub genes associated with physiological and biological traits. After filtering lowly expressed genes, expression values were imported to construct co-expression modules using the blockwiseModules function with a soft threshold power 10 (Fig. S2). Fig. S3 shows the cluster dendrogram of the genes after normalization of the validation dataset. As shown in Fig. 4A, we identified significantly associated modules with physiological or biological traits in NAFLD. ME14 (Module Eigen 14) module was significantly associated with the NAFLD trait. Accordingly, the correlation of the ME14 module with gene significance resulted in significance correlation of 0.83 (Fig. 4B). Based on this analysis, the co-expression network of ME14 module was further analyzed. We selected the genes with gene significance (GS) > 0.8 and module membership (MM) > 0.9 to construct the network regulatory map, and then highlighted gene significance associated with NAFLD trait (Fig. 4C). Pathway analysis of the significant genes has been carried out and shown in Fig. 4D. Most commonly enriched pathways include cytokine- cytokine receptor interaction pathway, JAK-STAT signaling pathway, PI3 K-Akt signaling pathway and pathways in cancer (Table S2).

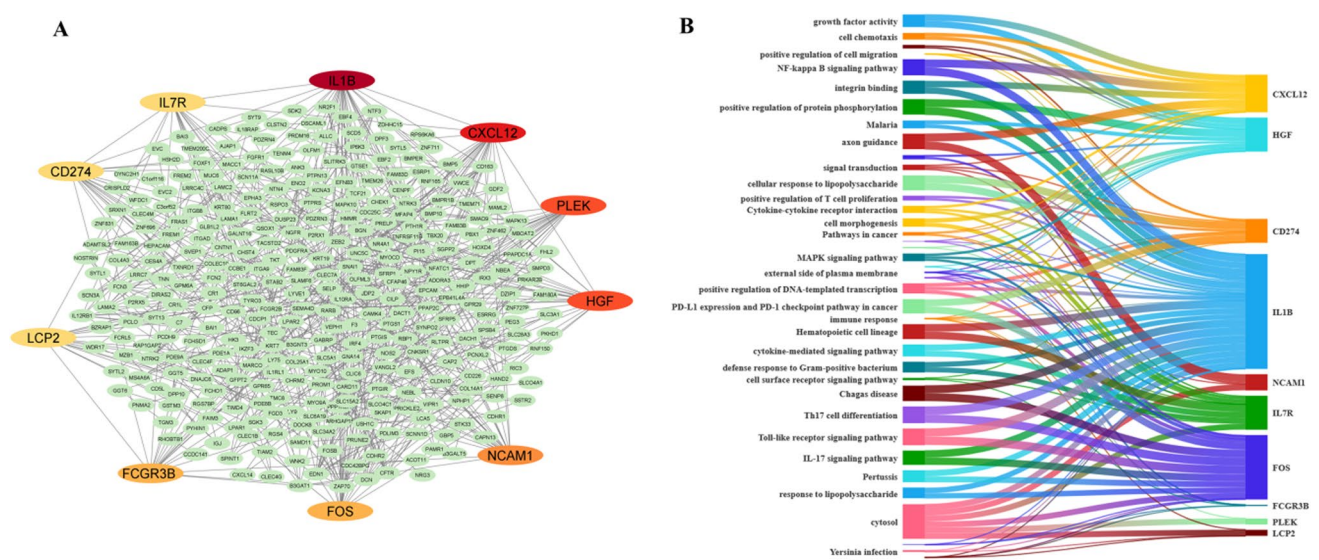


Fig. 3 Hub gene analysis. **A** Protein-proteins interaction network and colored nodes represent the highlighted top 10 hub genes. **B** Enriched terms of the hub genes

Table 1 Topology characteristics of hub genes

| Gene | Degree centrality | Average shortest path length | Betweenness centrality | Closeness centrality | Clustering coefficient | Neighborhood connectivity | Radiality | Stress |
|--------|-------------------|------------------------------|------------------------|----------------------|------------------------|---------------------------|-----------|---------|
| IL1B | 38 | 2.80 | 0.23 | 0.36 | 0.14 | 11.18 | 0.95 | 107,534 |
| CXCL12 | 25 | 2.94 | 0.09 | 0.33 | 0.20 | 13.16 | 0.94 | 49,004 |
| PLEK | 23 | 3.56 | 0.09 | 0.28 | 0.10 | 6.34 | 0.93 | 40,948 |
| HGF | 23 | 3.12 | 0.06 | 0.32 | 0.20 | 12.43 | 0.94 | 46,164 |
| NCAM1 | 22 | 3.02 | 0.12 | 0.33 | 0.23 | 13.32 | 0.94 | 52,216 |
| FCGR3B | 20 | 3.21 | 0.03 | 0.31 | 0.31 | 14 | 0.94 | 24,866 |
| FOS | 20 | 3.23 | 0.05 | 0.31 | 0.12 | 8.95 | 0.94 | 25,044 |
| LCP2 | 20 | 3.49 | 0.02 | 0.29 | 0.21 | 10.4 | 0.93 | 23,244 |
| CD274 | 20 | 3.16 | 0.03 | 0.32 | 0.34 | 15.05 | 0.94 | 21,042 |
| IL7R | 19 | 3.24 | 0.02 | 0.31 | 0.35 | 14.16 | 0.94 | 16,654 |

4 Discussion

Dysregulated metabolic pathways often act as a risk factor for cancer. The global incidence rate of MAFLD-related HCC has increased, distinct from other HCC etiologies, particularly in the absence of cirrhosis [62, 63]. Based on this research, our study involved the use of transcriptome datasets, which were obtained from the NCBI GEO database. The overview of the study is shown in Fig. 1. After preprocessing, the significant source of variation between the samples was visualized (Fig. 2A, B); these results are reflected in the number of DEGs obtained when comparing the samples from different groups as described by Law et al. [64, 65]. Volcano plot depicts the similarity of gene expression pattern from the MDS plot where the number of DEGs is more in the Fig. 2D, and the variance between the samples is higher in Fig. 2B when compared to Fig. 2C. Intersecting the DEGs resulted from both the dataset was 442 intersecting commonly expressed genes (30 up-regulated and 412 down-regulated) in comparison between the two transcriptome datasets (as shown in Fig. 2E). To evaluate the physical and functional associations among the DEGs, we utilized the PPI network using the STRING database [66]. PPI network was constructed using 442 DEGs as nodes, and genes without interactions were removed from the network, as shown in Fig. 3A. Hub genes were identified based on their topological characteristics; results are presented in Table 1. It includes average shortest path length, betweenness centrality, closeness centrality,

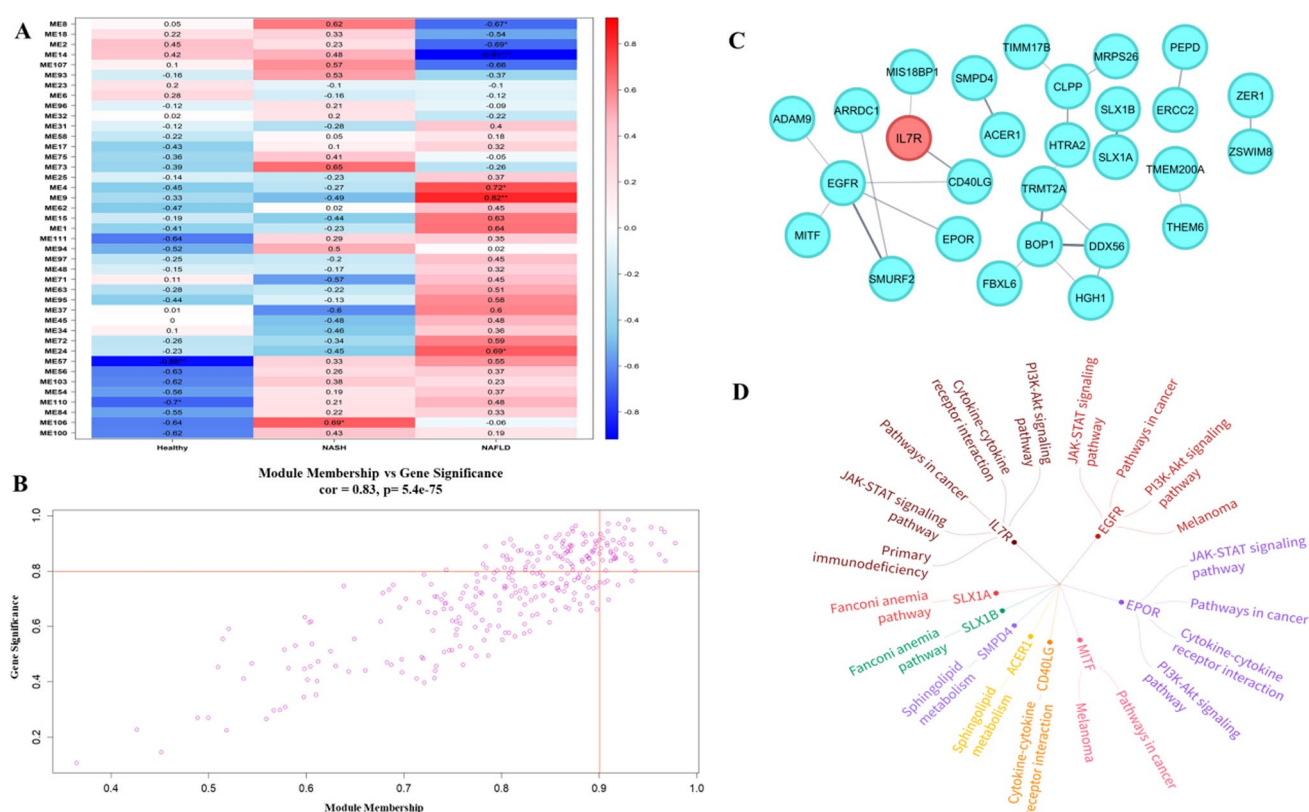


Fig. 4 WGCNA module identification. **A** Heat map of module-trait association. **B** Correlation of module (ME14) with gene significance. **C** Gene interaction network analysis highlighting significant genes. **D** Enriched pathways of significant genes

clustering coefficient, neighborhood connectivity, radiality, and stress. This study also utilized functional enrichment analysis to identify overrepresented biological pathways in the hub genes [67, 68]. The resulting hub genes have been reported to be involved in the inflammation, immune response, cell signalling, proliferation, cell adhesion and interaction, immune evasion and transcriptional regulations in HCC development, as highlighted in Fig. 3A [69–71]. Also, these results are reflected in Fig. 3B, which illustrates that most of the hub genes are enriched in the immune response, regulation of plasma membrane; integrin binding, cytokine signalling pathways and regulation of immune cells.

Subsequently, we approached WGCNA with an independent dataset to identify highly synergistic genomes and possible markers through an assessment of the interrelationship between genomes and their relationship to phenomena. For constructing co-expression modules we implemented blockwiseModules function with default settings, assigning the soft threshold power 10 and TOM Type is unsigned. By evaluating the interaction between each module and the molecular mechanism of cancer, the most prominent module ME14 was selected (Fig. 4A). The correlation of the module membership with gene significance is 0.83 (Fig. 4B). Gene interaction network was constructed based on the weighted genes in ME14 module (Fig. 4C). Enrichment pathway analysis of the significant genes has also been shown in Fig. 4D. From both the networks, it is evidently shown that most significant genes are cytokines or cytokine receptors which are majorly involved in the immune system regulation (Figs. 3A and 4C). The pathway analysis also reflects the functionally enriched cytokine-cytokine receptor signaling cascades. Consequently, based on the network analysis resulted from PPI network (Fig. 3A) and WGCNA module (Fig. 4C), the IL7R gene was found to be involved in both the analysis related to traits and as well as hub gene from the transcriptome analysis. As shown in Figs. 3B and 4D the driver gene IL7R, a cytokine receptor has been majorly involved in the cytokine-cytokine receptor signaling. Enrichment analysis shows that it is IL7R is majorly related to cytokine signaling functions (Table S1). The JAK/STAT signaling pathway, particularly via IL7R, is a crucial link between MAFLD and HCC because it promotes tumorigenesis, cell survival, and proliferation, leading to malignant transformation. The JAK/STAT pathway is particularly relevant in HCC due to its direct involvement in oncogenic processes driven by chronic inflammation and metabolic dysregulation [72, 73]. Its activation mechanisms through cytokines are a critical target for therapeutic intervention compared to the more general roles of PI3K and MAPK pathways [74]. The activation of the IL7R-JAK-STAT pathway is linked to alterations in the tumor microenvironment, affecting immune cell

infiltration and activity [75]. The immune infiltration analysis also shows that IL7R gene is negatively correlated with the tumor purity in LIHC samples in the immune cells CD8 + and CD4 + T cells (Fig. 5A). From KEGG analysis, IL7R interacts with JAK1 and JAK3 upon binding, activating JAK kinases and regulating gene expression [76]. Further validating the IL7R as a driver gene we utilized a logistic regression machine learning model to predict the malignancy of HCC based on the gene expression and sample pathology. It predicts the probability of observations belonging to different categorical outcomes and generalize linear model that relaxes assumptions, making it useful for non-continuous outcomes [77, 78]. We obtained an accuracy of 92% in predicting the HCC outcome in logistic regression machine learning model (Fig. 5C). Further, the model was evaluated for its performance through a precision-recall (PR) curve. From Fig. S4, we obtained an area under the curve (AUC) of 0.933 in the PR-curve analysis, which demonstrates that the model has a very high performance in distinguishing the positive and negative cases. In the same way, highly effective at identifying the positive instances while minimizing the false positives.

In the event of liver injury, HSCs are activated, leading to the transition from NAFLD to HCC, involving changes in the immune microenvironment, with increased infiltration in NAFLD and reduced cytotoxic T cells in HCC [79–82]. HSCs regulate hepatic homeostasis, maturation, and differentiation and are also involved in hepatic inflammation by releasing inflammatory cytokines and chemokines that interact with other liver cells [83]. IL-7, a pro-inflammatory cytokine released by the immune cells, binds to IL7R and involves the disease progression as shown in Fig. 5D [84–86]. Further, the IL-7/IL-7R axis activates the JAK/STAT pathway, specifically STAT3, which influences the HSC behavior [87, 88]. This axis is crucial in promoting T-cell development, differentiation, proliferation, and immune system reconstitution after depletion [89–93]. Polymorphisms in the IL7R gene are linked to liver stiffness, affecting the T-cell homeostasis and immune response in disease progression [94–96]. Immune infiltration analysis confirms T cell positive regulation in the HCC environment is correlated with enrichment analysis (ref Figs. 3B and 5A). IL7R gene expression changes influence macrophage interactions, epithelial to mesenchymal transition (EMT) promotion, T cell proliferation, and survival [97–99]. The immunosuppressive environment in HCC is characterized by reduced cytotoxic T cells and increased regulatory T cells, facilitating immune evasion and tumor progression. High IL7R expression in activated CD8 + T cells improves HCC patient survival [100–102]. In relation to that, our survival analysis also reflects the high-expression group has more survival than the low-expression groups as shown in Fig. 5B. Based on this evidence, activation of hepatic stellate cells triggers the IL-7/IL-7R axis, potentially reversing activation for disease management. To summarize, metabolic disorders cause lipotoxicity, leading to inflammatory responses in liver tissues. HSCs accumulate, triggering pro-inflammatory and

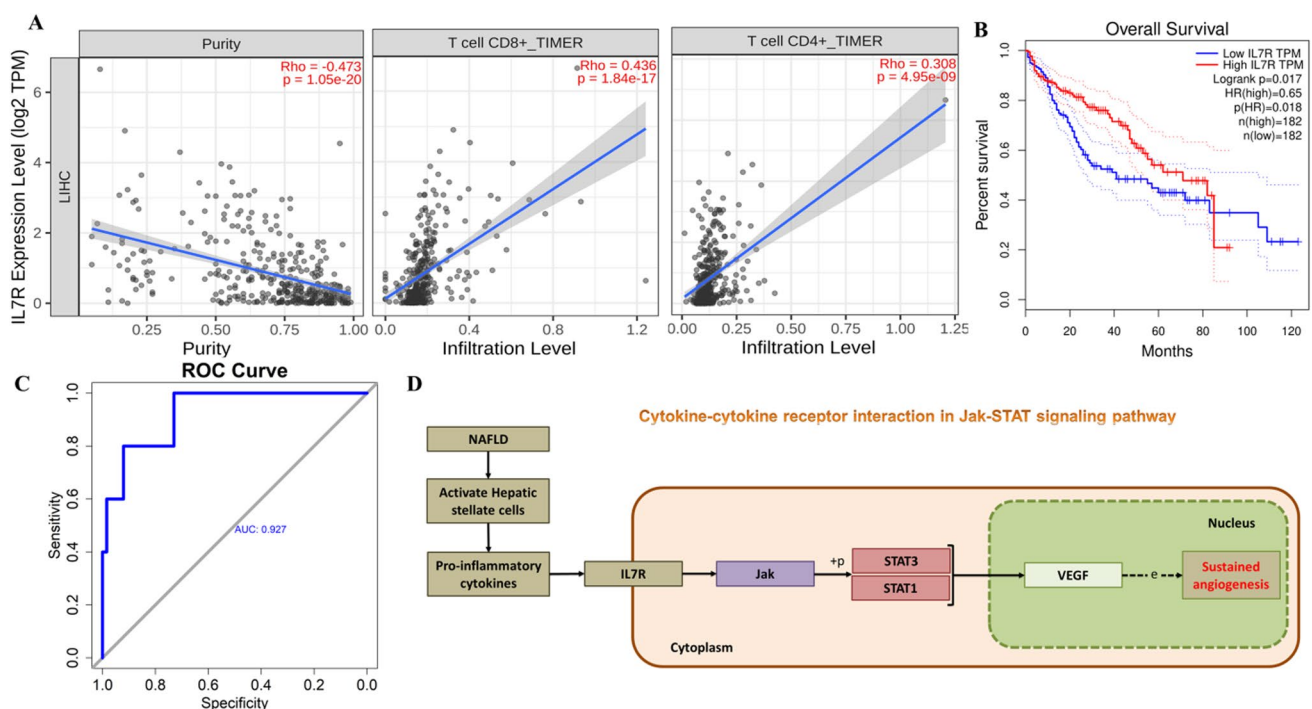


Fig. 5 Validation of the hub gene. **A** Immune infiltration of T cells. **B** Survival analysis. **C** ROC curve for validating IL7R expression. **D** Mechanism underlying the transition of fatty liver disease to HCC

pro-fibrogenic functions. IL7 binds to IL7R, triggering EMT. This process activates the transition of fatty liver disease to liver cancer majorly HCC involving immune microenvironment changes and promoting sustained angiogenesis. Targeting IL-7/IL-7R could offer a therapeutic approach for fibrosis-related liver diseases.

5 Conclusion

Metabolic dysfunction is a major factor in HCC invasion and managing inflammatory responses. We utilized transcriptome HCC datasets and identified the top ten hub genes from the PPI network. In addition to that, WGCNA was studied, and a gene interaction network was constructed to find trait-based gene significance. IL7R was identified as a driver gene for MAFLD traits in HCC. Tumors often exploit immune pathways for survival; the underlying mechanism behind the disease progression is the accretion of hepatic stellate cells, activating pro-inflammatory and pro-fibrogenic functions. The pro-inflammatory cytokines bind to the IL7R and trigger the JAK/STAT pathway, promoting cancer progression. Mitigating the IL-7/IL-7R axis could enhance immune surveillance and provide a therapeutic approach to liver disease management. A machine learning model predicts HCC malignancy based on IL7R gene expression, offering valuable insights into treatment options and early cancer management. Moreover, it can also be a perception for the diagnosis of non-cirrhotic HCC, where the development of HCC occurs in the absence of cirrhosis. Therefore, investigating the targets of IL7R in this context has the potential to provide valuable insights into treatment options for at-risk patients. It can also contribute to early cancer management in patients with underlying metabolic disorders. Targeting this pathway could offer new strategies for treating HCC by addressing the underlying inflammatory processes contributing to tumor progression. Further research is needed to elucidate how various cytokines contribute to aberrant activation of this pathway in different HCC subtypes, particularly in chronic liver diseases like NAFLD and NASH. Also, understanding these interactions could reveal how JAK/STAT mediates tumor progression and therapy resistance, especially in cases where multiple pathways are activated simultaneously. The limitation of the study is that it provides theoretical validation but requires further experimental validation. It offers insight into the pathogenesis of MAFLD progression from steatosis to HCC in non-cirrhotic patients. The MAFLD-HCC epidemic has resulted in a worldwide healthcare problem. Current therapy approaches do not consider etiology, resulting in underdiagnosis. Existing therapies for MAFLD-HCC are safe. However, randomized controlled studies are required to clarify therapy efficacy and stratify individuals.

6 Limitations

The study on MAFLD progression from steatosis to hepatocellular carcinoma (HCC) in non-cirrhotic patients has limitations, including a reliance on theoretical validation and the need for extensive experimental validation. The ongoing MAFLD-HCC epidemic is a global concern, with current therapeutic approaches often failing to account for underlying etiology, leading to under diagnosis and suboptimal treatment outcomes. Existing therapies have shown safety, but randomized controlled trials are needed to evaluate their efficacy and identify potential side effects. RCTs also help stratify patients based on clinical and genetic parameters, facilitating personalized therapeutic approaches. A comprehensive approach involving theoretical and experimental research, including clinical trials is needed to develop more effective and targeted treatments for MAFLD-HCC, addressing knowledge gaps and improving patient outcomes.

Acknowledgements The authors thank the Vellore Institute of Technology management for providing the resources to complete this study.

Author contributions All authors participated in the writing of the manuscript. Data collection, Data analysis, Statistical analysis and drafting the main text, figures and tables was carried out by P.A. Review and revising the article was carried out by A.E.I.

Funding Open access funding provided by Vellore Institute of Technology. No funding was received to conduct this study.

Data availability The GEO database from NCBI (Gene Expression Omnibus database, <https://www.ncbi.nlm.nih.gov/geo/>) was used to access the GSE140462, GSE184733 and GSE115193 datasets.

Declarations

Ethics approval and consent to participate Ethical approval is not required. No human or animal participants were involved in this study. All the data utilized for this study were retrieved from public repositories.

Competing interests The authors declare that they have no competing interests.

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

References

1. Zha X, Gao Z, Li M, Xia X, Mao Z, Wang S. Insight into the regulatory mechanism of m6A modification: from MAFLD to hepatocellular carcinoma. *Biomed Pharmacother*. 2024;177:116966. <https://doi.org/10.1016/j.biopha.2024.116966>.
2. Lin YP, Wang PM, Chuang CH, Yong CC, Liu YW, Huang PY, et al. Metabolic risks are increasing in Non-B Non-C early-stage hepatocellular carcinoma: a 10-year follow-up study. *Front Oncol*. 2022;12:1–10. <https://doi.org/10.3389/fonc.2022.816472>.
3. Boccata A, Andreotto L, D'Ardes D, Cocco G, Rossi I, Vicari S, et al. From NAFLD to MAFLD: definition. *Pathophysiol Basis Cardiovasc Implic Biomed*. 2023;11:1–15. <https://doi.org/10.3390/biomedicines11030883>.
4. Xue R, Fan J. Brief introduction of an international expert consensus statement: a new definition of metabolic associated fatty liver disease. *J Clin Hepatol*. 2020;36:1224–7. <https://doi.org/10.3969/j.issn.1001-5256.2020.06.007>.
5. Angulo P. Nonalcoholic fatty liver disease. *N Engl J Med*. 2002;346:1221–31. <https://doi.org/10.1056/NEJMra011775>.
6. Rinella ME. Nonalcoholic fatty liver disease a systematic review. *JAMA*. 2015;313:2263–73. <https://doi.org/10.1001/jama.2015.5370>.
7. Estes C, Anstee QM, Arias-Loste MT, Bantel H, Bellentani S, Caballeria J, et al. Modeling NAFLD disease burden in China, France, Germany, Italy, Japan, Spain, United Kingdom, and United States for the period 2016–2030. *J Hepatol*. 2018;69:896–904. <https://doi.org/10.1016/j.jhep.2018.05.036>.
8. Grgurevic I, Bozin T, Mikus M, Kukla M, O'beirne J. Hepatocellular carcinoma in non-alcoholic fatty liver disease: from epidemiology to diagnostic approach. *Cancers*. 2021. <https://doi.org/10.3390/cancers13225844>.
9. Pinyopornpanish K, Khoudari G, Saleh MA, Angkurawaranon C, Pinyopornpanish K, Mansoor E, et al. Hepatocellular carcinoma in non-alcoholic fatty liver disease with or without cirrhosis: a population-based study. *BMC Gastroenterol*. 2021;21:1–7. <https://doi.org/10.1186/s12876-021-01978-0>.
10. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA*. 2023. <https://doi.org/10.3322/caac.21763>.
11. Wang Y, Fleishman JS, Li T, Li Y, Ren Z, Chen J, et al. Pharmacological therapy of metabolic dysfunction-associated steatotic liver disease-driven hepatocellular carcinoma. *Front Pharmacol*. 2023;14:1–15. <https://doi.org/10.3389/fphar.2023.1336216>.
12. Panahi B. Global transcriptome analysis identifies critical functional modules associated with multiple abiotic stress responses in microalgae *Chromochloris zofingiensis*. *PLoS ONE*. 2024;19:1–19. <https://doi.org/10.1371/journal.pone.0307248>.
13. Norero B, Dufour JF. Should we undertake surveillance for HCC in patients with MAFLD? *Ther Adv Endocrinol Metab*. 2023;14:1–8. <https://doi.org/10.1177/20420188231160389>.
14. Sun Y, Zhang H, Li Y, Han J. Abnormal metabolism in hepatic stellate cells: pandora's box of MAFLD related hepatocellular carcinoma. *Biochim et Biophys Acta (BBA) Rev Cancer*. 2024. <https://doi.org/10.1016/j.bbcan.2024.189086>.
15. Crane H, Gofton C, Sharma A, George J. MAFLD: an optimal framework for understanding liver cancer phenotypes. *J Gastroenterol*. 2023;58:947–64. <https://doi.org/10.1007/s00535-023-02021-7>.
16. Chavez-Tapia NC, Murúa-Beltrán Gall S, Ordoñez-Vázquez AL, Nuño-Lambarri N, Vidal-Cevallos P, Uribe M. Understanding the role of metabolic syndrome as a risk factor for hepatocellular carcinoma. *JHC*. 2022;9:583–93. <https://doi.org/10.2147/jhc.s283840>.
17. Llovet JM, Kelley RK, Villanueva A, Singal AG, Pikarsky E, Roayaie S, et al. Hepatocellular carcinoma. *Nat Rev Dis Primers*. 2021;7:6. <https://doi.org/10.1038/s41572-020-00240-3>.
18. Wang H, Zhao Y, Pan Y, Yang A, Li C, Wang S, et al. Inhibition of phospholipase D1 ameliorates hepatocyte steatosis and non-alcoholic fatty liver disease. *JHEP Rep*. 2023;5:100726. <https://doi.org/10.1016/j.jhepr.2023.100726>.
19. Petrelli F, Manara M, Colombo S, De Santi G, Ghidini M, Mariani M, et al. Hepatocellular carcinoma in patients with nonalcoholic fatty liver disease: a systematic review and meta-analysis. *Neoplasia*. 2022;30:100809. <https://doi.org/10.1016/j.neo.2022.100809>.
20. Wei S, Hao Y, Dong X, Huang J, Huang K, Xie Y, et al. The relationship between metabolic dysfunction-associated fatty liver disease and the incidence rate of extrahepatic cancer. *Front Endocrinol*. 2023;14:1–8. <https://doi.org/10.3389/fendo.2023.985858>.
21. Chavez-Tapia NC, Mendez-Sanchez N, Uribe M. Role of nonalcoholic fatty liver disease in hepatocellular carcinoma. *Ann Hepatol*. 2009;8:34–9. [https://doi.org/10.1016/s1665-2681\(19\)31824-1](https://doi.org/10.1016/s1665-2681(19)31824-1).
22. Tong L, Chen Z, Li Y, Wang X, Yang C, Li Y, et al. Transketolase promotes MAFLD by limiting inosine-induced mitochondrial activity. *Cell Metab*. 2024;36:1013–1029.e5. <https://doi.org/10.1016/j.cmet.2024.03.003>.
23. Schulz PO, Ferreira FG, De Nascimento MFA, Vieira A, Ribeiro MA, David AI, et al. Association of nonalcoholic fatty liver disease and liver cancer. *WJG*. 2015. <https://doi.org/10.3748/wjg.v21.i3.913>.
24. Fabris L, Campello E, Cadamuro M, Simioni P. The evil relationship between liver fibrosis and cardiovascular disease in metabolic dysfunction-associated fatty liver disease (MAFLD): looking for the culprit. *Biochim et Biophys Acta Mol Basis Dis*. 2024;1870:166763. <https://doi.org/10.1016/j.bbdis.2023.166763>.
25. Chaturvedi S, Tripathi D, Vikram NK, Madhusudan KS, Pandey RM, Bhatia N. Association of nutrient intake with non-alcoholic fatty liver disease and liver steatosis in adult Indian population—a case control study. *Human Nutr Metab*. 2023;32:200188. <https://doi.org/10.1016/j.hnm.2023.200188>.

26. AmeliMojarad M, AmeliMojarad M, Cui X. Discovering the lipid metabolism-related hub genes of HCC-treated samples with PPAR α agonist through weighted correlation network analysis. *Sci Rep*. 2024;14:1–12. <https://doi.org/10.1038/s41598-024-69998-w>.
27. Paternostro R, Sieghart W, Trauner M, Pinter M. Cancer and hepatic steatosis. *ESMO Open*. 2021;6:100185. <https://doi.org/10.1016/j.esmoop.2021.100185>.
28. Chiang JYL, Li T. Pathogenesis of fatty liver diseases and hepatocellular carcinoma. *Liver Res*. 2022;6:201–2. <https://doi.org/10.1016/j.livres.2022.12.001>.
29. Ryu T, Kim K, Choi SE, Chung KPS, Jeong W-I. New insights in the pathogenesis of alcohol-related liver disease: the metabolic, immunologic, and neurologic pathways. *Liver Res*. 2022;7:1–8. <https://doi.org/10.1016/j.livres.2022.09.004>.
30. Mittal S, El-Serag HB, Sada YH, Kanwal F, Duan Z, Temple S, et al. Hepatocellular carcinoma in the absence of cirrhosis in United States veterans is associated with nonalcoholic fatty liver disease. *Clin Gastroenterol Hepatol*. 2016;14:124–131.e1. <https://doi.org/10.1016/j.cgh.2015.07.019>.
31. Efreim IC, Moța M, Vladu IM, Mitrea A, Clenciu D, Timofticiuc DCP, et al. A study of biomarkers associated with metabolic dysfunction-associated fatty liver disease in patients with type 2 diabetes. *Diagnostics*. 2022. <https://doi.org/10.3390/diagnostics12102426>.
32. Gofton C, Upendran Y, Zheng MH, George J. MAFLD: how is it different from NAFLD? *Clin Mol Hepatol*. 2023;29:S17–31. <https://doi.org/10.3350/cmh.2022.0367>.
33. Sangro P, de la Torre AM, Sangro B, D'Avola D. Metabolic dysfunction-associated fatty liver disease (MAFLD): an update of the recent advances in pharmacological treatment. *J Physiol Biochem*. 2023;79:869–79. <https://doi.org/10.1007/s13105-023-00954-4>.
34. Gao F, Zheng KI, Yan HD, Sun QF, Pan KH, Wang TY, et al. Association and interaction between serum interleukin-6 levels and metabolic dysfunction-associated fatty liver disease in patients with severe coronavirus disease 2019. *Front Endocrinol*. 2021;12:1–9. <https://doi.org/10.3389/fendo.2021.604100>.
35. Guo Q, Fan Y, Xie M, Wang Q, Zou Z, Gao G, et al. Exploring the transcriptomic landscape of moyamoya disease and systemic lupus erythematosus: insights into crosstalk genes and immune relationships. *Front Immunol*. 2024. <https://doi.org/10.3389/fimmu.2024.1456392>.
36. Sánchez-Baizán N, Ribas L, Piferrer F. Improved biomarker discovery through a plot twist in transcriptomic data analysis. *BMC Biol*. 2022;20:1–26. <https://doi.org/10.1186/s12915-022-01398-w>.
37. Giulietti M, Occhipinti G, Righetti A, Bracci M, Conti A, Ruzzo A, et al. Emerging biomarkers in bladder cancer identified by network analysis of transcriptomic data. *Front Oncol*. 2018;8:1–8. <https://doi.org/10.3389/fonc.2018.00450>.
38. Clough E, Barrett T. The gene expression omnibus database. *Methods Mol Biol*. 2016;1418:93–110. https://doi.org/10.1007/978-1-4939-3578-9_5.
39. Hall Z, Chiarugi D, Charidemou E, Leslie J, Scott E, Pellegrinet L, et al. Lipid remodeling in hepatocyte proliferation and hepatocellular carcinoma. *Hepatology*. 2021;73:1028–44. <https://doi.org/10.1002/hep.31391>.
40. Liu Y, Al-Adra DP, Lan R, Jung G, Li H, Yeh MM, et al. RNA sequencing analysis of hepatocellular carcinoma identified oxidative phosphorylation as a major pathologic feature. *Hepatol Commun*. 2022;6:2170–81. <https://doi.org/10.1002/hep4.1945>.
41. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*. 2009;26:139–40. <https://doi.org/10.1093/bioinformatics/btp616>.
42. Law CW, Chen Y, Shi W, Smyth GK. Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol*. 2014;15:1–17. <https://doi.org/10.1186/gb-2014-15-2-r29>.
43. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res*. 2019;47:D607–13. <https://doi.org/10.1093/nar/gky1131>.
44. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T, et al. Cytoscape: a software environment for integrated models. *Genome Res*. 1971;13:426. <https://doi.org/10.1101/gr.1239303.metabolite>.
45. Doncheva NT, Morris JH, Gorodkin J, Jensen LJ. Cytoscape stringapp: network analysis and visualization of proteomics data. *J Proteome Res*. 2019;18:623–32. <https://doi.org/10.1021/acs.jproteome.8b00702>.
46. Bader GD, Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinform*. 2001;29:137–40.
47. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*. 2014;8:1–7. <https://doi.org/10.1186/1752-0509-8-S4-S11>.
48. Grindrod P, Kibble M. Review of uses of network and graph theory concepts within proteomics. *Expert Rev Proteomics*. 2004;1:229–38. <https://doi.org/10.1586/14789450.1.2.229>.
49. Metcalf L, Casey W. Cybersecurity and applied mathematics. Amsterdam: Elsevier; 2016. <https://doi.org/10.1016/B978-0-12-804452-0.00005-1>.
50. Nafis S, Kalaiaresan P, Singh RKB, Husain M, Bamezai RNK. Apoptosis regulatory protein-protein interaction demonstrates hierarchical scale-free fractal network. *Brief Bioinform*. 2014;16:675–99. <https://doi.org/10.1093/bib/bbu036>.
51. AbdulHameed MDM, Ippolito DL, Stallings JD, Wallqvist A. Mining kidney toxicogenomic data by using gene co-expression modules. *BMC Genomics*. 2016;17:1–17. <https://doi.org/10.1186/s12864-016-3143-y>.
52. Febbraio MA, Reibe S, Ooi GJ, Watt MJ, Karin M. Preclinical models for studying NASH-driven HCC: how useful are they? *Cell Metab*. 2019;29:18–26. <https://doi.org/10.1016/j.cmet.2018.10.012>.
53. Hou J, Ye X, Feng W, Zhang Q, Han Y, Liu Y, et al. Distance correlation application to gene co-expression network analysis. *BMC Bioinform*. 2022;23:1–24. <https://doi.org/10.1186/s12859-022-04609-x>.
54. Sherman BT, Hao M, Qiu J, Jiao X, Baseler MW, Lane HC, et al. DAVID: a web server for functional enrichment analysis and functional annotation of gene lists (2021 update). *Nucleic Acids Res*. 2022;50:W216–21. <https://doi.org/10.1093/nar/gkac194>.
55. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B Stat Methodol*. 1995;57:289–300. <https://doi.org/10.1111/j.2517-6161.1995.tb02031.x>.
56. Lian Q, Wang S, Zhang G, Wang D, Luo G, Tang J, et al. HCCDB: a database of hepatocellular carcinoma expression Atlas. *Genom Proteom Bioinform*. 2018;16:269–75. <https://doi.org/10.1016/j.gpb.2018.07.003>.

57. Tripepi G, Jager KJ, Dekker FW, Zoccali C. Linear and logistic regression analysis. *Kidney Int.* 2008;73:806–10. <https://doi.org/10.1038/sj.ki.5002787>.
58. Glazko G, Mushegian A. Measuring gene expression divergence: the distance to keep. *Biol Direct.* 2010;5:1–10. <https://doi.org/10.1186/1745-6150-5-51>.
59. Zhou Y, Sharpee TO. Hyperbolic geometry of gene expression. *iScience.* 2021. <https://doi.org/10.1016/j.isci.2021.102225>.
60. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43:e47. <https://doi.org/10.1093/nar/gkv007>.
61. Fonseka P, Pathan M, Chitti SV, Kang T, Mathivanan S. FunRich enables enrichment analysis of OMICs datasets. *J Mol Biol.* 2021;433:166747. <https://doi.org/10.1016/j.jmb.2020.166747>.
62. Barakat EM, Montasser IF, Zaky DZ, Abdelrazik YA, Farid HM, El DA, et al. Outcome of MAFLD-related HCC in Egyptian patients: a single center study. *Egypt Liver J.* 2024. <https://doi.org/10.1186/s43066-024-00337-4>.
63. Demirtaş CÖ, Tolu T, Keklikkiran Ç, Özdoğan OC, Gündüz F. Hepatocellular carcinoma in non-cirrhotic liver arises with a more advanced tumoral appearance: a single-center cohort study. *Turk J Gastroenterol.* 2021;32:685–93. <https://doi.org/10.5152/tjg.2021.20677>.
64. Law CW, Alhamdoosh M, Su S, Dong X, Tian L, Smyth GK, et al. RNA-seq analysis is easy as 1-2-3 with limma, Glimma and edgeR. *F1000Res.* 2018. <https://doi.org/10.12688/f1000research.9005.3>.
65. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of RNA-seq data. *Genome Biol.* 2010;11:1–9.
66. Athanasios A, Charalampos V, Vasileios T, Ashraf G. Protein-protein interaction (PPI) network: recent advances in drug discovery. *Curr Drug Metab.* 2017;18:5–10. <https://doi.org/10.2174/138920021801170119204832>.
67. Chicco D, Agapito G. Nine quick tips for pathway enrichment analysis. *PLoS Comput Biol.* 2022;18:1–15. <https://doi.org/10.1371/journal.pcbi.1010348>.
68. Ashburner M, Ball CA, Blake JA, Butler H, Cherry JM, Eppig JT, et al. Creating the gene ontology resource: design and implementation. *Genome Res.* 2001;11:1425–33. <https://doi.org/10.1101/gr.180801>.
69. Chen W, Jiang J, Wang PP, Gong L, Chen J, Du W, et al. Identifying hepatocellular carcinoma driver genes by integrative pathway crosstalk and protein interaction network. *DNA Cell Biol.* 2019;38:1112–24. <https://doi.org/10.1089/dna.2019.4869>.
70. Jiang M, Chen Y, Zhang Y, Chen L, Zhang N, Huang T, et al. Identification of hepatocellular carcinoma related genes with k-th shortest paths in a protein-protein interaction network. *Mol BioSyst.* 2013;9:2720–8. <https://doi.org/10.1039/c3mb70089e>.
71. Song H, Ding N, Li S, Liao J, Xie A, Yu Y, et al. Identification of hub genes associated with hepatocellular carcinoma using robust rank aggregation combined with weighted gene co-expression network analysis. *Front Genet.* 2020;11:1–14. <https://doi.org/10.3389/fgene.2020.00895>.
72. Park H, Lee S, Lee J, Moon H, Ro SW. Exploring the JAK/STAT signaling pathway in hepatocellular carcinoma: unraveling signaling complexity and therapeutic implications. *IJMS.* 2023. <https://doi.org/10.3390/ijms241813764>.
73. Banerjee A, Das D, Maji BK. Non-alcoholic steatohepatitis and hepatocellular carcinoma are influenced by oxidative stress, inflammation, and apoptosis via activating the AMPK- caspase axis and JAK / STAT pathway. *JCH.* 2024. https://doi.org/10.25259/JCH_7_2023.
74. Hin Tang JJ, Hao Thng DK, Lim JJ, Toh TB. JAK/STAT signaling in hepatocellular carcinoma. *Hepat Oncol.* 2020. <https://doi.org/10.2217/hep-2020-0001>.
75. Lodewijckx I, Mentens N, Jacobs K, Cools J. Oncogenic cooperation between IL7R-JAK-STAT pathway mutations. *HemaSphere.* 2021;5:e637. <https://doi.org/10.1097/HS9.0000000000000637>.
76. Huang C, Jiang X, Huang Y, Zhao L, Li P, Liu F. Identifying dendritic cell-related genes through a co-expression network to construct a 12-gene risk-scoring model for predicting hepatocellular carcinoma prognosis. *Front Mol Biosci.* 2021;8:1–16. <https://doi.org/10.3389/fmolb.2021.636991>.
77. Zaidi A, Al Luhayb ASM. Two statistical approaches to justify the use of the logistic function in binary logistic regression. *Math Probl Eng.* 2023. <https://doi.org/10.1155/2023/5525675>.
78. Starbuck C. The fundamentals of people analytics with applications in R. Cham: Springer International Publishing; 2023. <https://doi.org/10.1007/978-3-031-28674-2>.
79. Weiskirchen R. Pharmacological intervention in hepatic stellate cell activation and hepatic fibrosis. *Front Pharmacol.* 2016. <https://doi.org/10.3389/fphar.2016.00033>.
80. Zisser A, Ipsen DH, Tveden-Nyborg P. Hepatic stellate cell activation and inactivation in nash-fibrosis—roles as putative treatment targets? *Biomedicines.* 2021;9:1–18. <https://doi.org/10.3390/biomedicines9040365>.
81. Lu Y, Li M, Zhou Q, Fang D, Wu R, Li Q, et al. Dynamic network biomarker analysis and system pharmacology methods to explore the therapeutic effects and targets of Xiaoyaosan against liver cirrhosis. *J Ethnopharmacol.* 2022;294:115324. <https://doi.org/10.1016/j.jep.2022.115324>.
82. Zeng F, Zhang Y, Han X, Zeng M, Gao Y, Weng J. Predicting non-alcoholic fatty liver disease progression and immune deregulations by specific gene expression patterns. *Front Immunol.* 2021;11:1–18. <https://doi.org/10.3389/fimmu.2020.609900>.
83. Gupta G, Khadem F, Uzonna JE. Role of hepatic stellate cell (HSC)-derived cytokines in hepatic inflammation and immunity. *Cytokine.* 2019;124:154542. <https://doi.org/10.1016/j.cyto.2018.09.004>.
84. Lodewijckx I, Cools J. Deregulation of the interleukin-7 signaling pathway in lymphoid malignancies. *Pharmaceuticals.* 2021. <https://doi.org/10.3390/ph14050443>.
85. Wang DH, Ye LH, Ning JY, Zhang XK, Lv TT, Li ZJ, et al. Single-cell sequencing and multiple machine learning algorithms to identify key T-cell differentiation gene for progression of NAFLD cirrhosis to hepatocellular carcinoma. *Front Mol Biosci.* 2024;11:1–14. <https://doi.org/10.3389/fmolb.2024.1301099>.
86. Ampuero S, Bahamonde G, Tempio F, Garmendia ML, Ruiz M, Pizarro R, et al. IL-7/IL7R axis dysfunction in adults with severe community-acquired pneumonia (CAP): a cross-sectional study. *Sci Rep.* 2022;12:13145. <https://doi.org/10.1038/s41598-022-13063-x>.
87. Webb K, Williams A, Melendez Q, Wooten C, Greene N, Lopez D, et al. PSUN300 the sex-specific correlation of IL7, high HbA1c and liver damage markers in African Americans. *J Endocrine Soc.* 2022;6:A290–1. <https://doi.org/10.1210/jendso/bvac150>.
88. Perito ER, Ajmera V, Bass NM, Rosenthal P, Lavine JE, Schwimmer JB, et al. Association between cytokines and liver histology in children with nonalcoholic fatty liver disease. *Hepatol Commun.* 2017;1:609–22. <https://doi.org/10.1002/hep4.1068>.

89. Ceausu A, Rodríguez-Gallego E, Peraire J, López-Dupla M, Domingo P, Viladés C, et al. IL-7/IL-7R gene variants impact circulating IL-7/IL-7R homeostasis and ART-associated immune recovery status. *Sci Rep*. 2019;9:1–10. <https://doi.org/10.1038/s41598-019-52025-8>.
90. He K, Liu S, Xia Y, Xu J, Liu F, Xiao J, et al. CXCL12 and IL7R as novel therapeutic targets for liver hepatocellular carcinoma are correlated with somatic mutations and the tumor immunological microenvironment. *Front Oncol*. 2020;10:1–13. <https://doi.org/10.3389/fonc.2020.574853>.
91. Rafaqat S, Gluscevic S, Mercantepe F, Rafaqat S, Klisic A. Interleukins: pathogenesis in non-alcoholic fatty liver disease. *Metabolites*. 2024. <https://doi.org/10.3390/metabo14030153>.
92. Barata JT, Durum SK, Seddon B. Flip the coin: IL-7 and IL-7R in health and disease. *Nat Immunol*. 2019;20:1584–93. <https://doi.org/10.1038/s41590-019-0479-x>.
93. Fernandes MB, Barata JT. IL-7 and IL-7R in health and disease: an update through COVID times. *Adv Biol Regul*. 2023. <https://doi.org/10.1016/j.jbior.2022.100940>.
94. Midorikawa Y, Tsutsumi S, Taniguchi H, Ishii M, Kobune Y, Kodama T, et al. Identification of genes associated with dedifferentiation of hepatocellular carcinoma with expression profiling analysis. *Jpn J Cancer Res*. 2002;93:636–43. <https://doi.org/10.1111/j.1349-7006.2002.tb01301.x>.
95. Zhang J, Wang L, Jiang M. Diagnostic value of sphingolipid metabolism-related genes CD37 and CXCL9 in nonalcoholic fatty liver disease. *Medicine*. 2024;103:E37185. <https://doi.org/10.1097/MD.00000000000037185>.
96. Jiménez-Sousa MÁ, Gómez-Moreno AZ, Pineda-Tenor D, Medrano LM, Sánchez-Ruano JJ, Fernández-Rodríguez A, et al. The IL7RA rs6897932 polymorphism is associated with progression of liver fibrosis in patients with chronic hepatitis C: repeated measurements design. *PLoS ONE*. 2018;13:1–13. <https://doi.org/10.1371/journal.pone.0197115>.
97. Lev A, Simon AJ, Barel O, Eyal E, Glick-Saar E, Nayshool O, et al. Reduced function and diversity of t cell repertoire and distinct clinical course in patients with IL7 mutation. *Front Immunol*. 2019;10:1–13. <https://doi.org/10.3389/fimmu.2019.01672>.
98. Kong F, Hu W, Zhou K, Wei X, Kou Y, You H, et al. Hepatitis B virus X protein promotes interleukin-7 receptor expression via NF-κB and Notch1 pathway to facilitate proliferation and migration of hepatitis B virus-related hepatoma cells. *J Exp Clin Cancer Res*. 2016;35:1–14. <https://doi.org/10.1186/s13046-016-0448-2>.
99. Herck MAV, Weyler J, Kwanten WJ, Dirinck EL, Winter BYD, Francque SM, et al. The differential roles of T-cells in non-alcoholic fatty liver disease and obesity. *Front Immunol*. 2019. <https://doi.org/10.3389/fimmu.2019.00082>.
100. Hu R, Zhang W, Han Z, Ma M, Huang Q, Lv M, et al. Identification of immune-related target and prognostic biomarkers in PBMC of hepatocellular carcinoma. *BMC Gastroenterol*. 2023;23:234. <https://doi.org/10.1186/s12876-023-02843-y>.
101. Zhuang W, Sun H, Zhang S, Zhou Y, Weng W, Wu B, et al. An immunogenomic signature for molecular classification in hepatocellular carcinoma. *Mol Ther Nucleic Acids*. 2021;25:105–15. <https://doi.org/10.1016/j.omtn.2021.06.024>.
102. To JC, Gao S, Li XX, Zhao Y, Keng VW. Sorafenib resistance contributed by IL7 and MAL2 in hepatocellular carcinoma can be overcome by autophagy-inducing stapled peptides. *Cancers*. 2023. <https://doi.org/10.3390/cancers15215280>.

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