TPM Vol. 32, No. S2, 2025
ISSN: 1972-6325
https://www.tpmap.org/

Open Access

IDENTIFICATION OF AGE-RELATED HUB GENES IN HEPATITIS B-ASSOCIATED HEPATOCELLULAR CARCINOMA THROUGH TRANSCRIPTOMIC ANALYSIS: INSIGHTS INTO PROGNOSTIC BIOMARKERS AND VACCINE DEVELOPMENT

SANDHYA ANANTH[1], DEBALINA SEN[1], AROCKIA ALEX[2]

[1,2]DEPARTMENT OF BIOCHEMISTRY, SAVEETHA MEDICAL COLLEGE, SAVEETHA INSTITUTE OF MEDICAL AND TECHNICAL SCIENCES, SAVEETHA UNIVERSITY, CHENNAI, TAMIL NADU, INDIA.

CORRESPONDING AUTHOR*:
AROCKIA ALEX (alexrock2230@gmail.com)

Abstract

Hepatocellular carcinoma (HCC) continues to pose a considerable worldwide health challenge, especially in areas with elevated Hepatitis B Virus (HBV) incidence. Notwithstanding progress in immunization and antiviral therapies, chronic HBV carriers continue to face danger. The genetic diversity of HCC requires the discovery of new prognostic biomarkers and therapeutic targets. This research investigates age-related hub genes in HBV-related hepatocellular carcinoma by transcriptome analysis. Gene expression data from GEO (GSE45267 and GSE38941) were examined to detect differentially expressed genes (DEGs) with GEO2R, then constructing a protein-protein interaction (PPI) network through STRING and Cytoscape. Functional enrichment and survival studies were conducted with GO, KEGG, UALCAN, GEPIA, and KM plotter. We discovered 1,322 differentially expressed genes (DEGs) in HBV samples and 249 overlapping DEGs in both young and elderly hepatocellular carcinoma (HCC) patients. PPI study identified significant hub genes: SYK, C3, CD44, IGLL5 (associated with HBV) and CDK1, UBE2C, BUB1B, CCNB1, CDC20, CCNA2, MAD2L1, BIRC5 (associated with HCC). These genes participate in immune response, metabolism, and tumour growth, with increased expression associated with worse prognosis. Our research elucidates essential molecular pathways linked with HBV-related HCC and aging-related tumour biology, pinpointing prospective biomarkers and therapeutic targets for precision medicine and vaccine innovation.

Keywords: Hepatocellular carcinoma, HBV, Age-related hub genes, Prognostic biomarkers, Transcriptome analysis, Vaccine development.

INTRODUCTION

Hepatocellular carcinoma (HCC) is a significant global health issue, being the sixth most often diagnosed cancer and the third primary cause of cancer-related mortality globally. Many HCC cases are linked to long-term infection with the Hepatitis B Virus (HBV) (Abdelhamed & El-Kassas, 2024; Asafo-Agyei & Samant, 2025; Llovet et al., 2021). This is especially true in places where HBV is common, like Asia and sub-Saharan Africa. Hepatitis B virus-associated hepatocellular carcinoma is influenced by a confluence of viral, genetic, and environmental variables, whereby chronic inflammation, hepatocyte necrosis, and fibrosis are pivotal in the advancement of the illness (Di Bisceglie, 2009; Rizzo et al., 2022; C. Shen et al., 2023). The incorporation of the HBV genome into host DNA results in genomic instability, deregulation of oncogenes and tumour suppressor genes, and the activation of pro-inflammatory and fibrotic pathways. The molecular modifications, together with the immune evasion tactics used by HBV, facilitate hepatocarcinogenesis and tumour advancement (Hai et al., 2014; M. Zhang et al., 2024; Zhao et al., 2016). Despite progress in HBV vaccines and antiviral treatments, the risk of HCC remains elevated in chronic HBV carriers,

highlighting the urgent need for dependable prognostic biomarkers and innovative therapeutic approaches. A major issue in hepatocellular carcinoma care is the absence of early diagnostic signs since symptoms often manifest at late stages when curative interventions like liver transplantation and surgical resection are no longer feasible (C. Shen et al., 2023). Current biomarkers, including alpha-fetoprotein (AFP), have inadequate specificity and sensitivity, rendering them unsuitable for precise early diagnosis. In this setting, transcriptome analysis has emerged as a potent method to reveal molecular markers linked to HCC development (Chan et al., 2024; Pan et al., 2020; Parikh et al., 2020). By analysing gene expression patterns, researchers may discern differentially expressed genes and important regulatory networks that influence disease development. Hub genes, which are centrally linked within gene interaction networks, are crucial in oncogenesis, immunological regulation, and treatment resistance (Rosati et al., 2024; Xue et al., 2020). Identifying these hub genes offers essential insights into prospective targets for the creation of predictive biomarkers and tailored treatment therapies. The ageing process is a crucial factor affecting HCC growth since the molecular and immunological characteristics of tumours change markedly across various age demographics (Hossen et al., 2024; Mi et al., 2020). Patients of younger age often display aggressive tumour phenotypes characterized by distinct genetic abnormalities, while older patients may have a more immune-suppressed tumour microenvironment attributable to age-related modifications in immune surveillance and chronic inflammation. The relationship between ageing and HBV-induced hepatocarcinogenesis is little comprehended, necessitating an examination of age-related molecular changes in HCC (Duan et al., 2024; Fane & Weeraratna, 2020). Comprehending the impact of age on gene expression patterns in HBV-associated HCC would facilitate the discovery of shared and unique molecular signatures across various age demographics, hence establishing the foundation for age-specific prognostic indicators and therapeutic targets (Atyah et al., 2018; Yip et al., 2017; Zoulim et al., 2024).

This research seeks to discover age-related hub genes in HBV-associated hepatocellular carcinoma via transcriptome analysis. We want to clarify the molecular processes behind hepatocarcinogenesis at various life stages by comparing gene expression patterns across different age groups. The discovery of these hub genes may augment early detection efforts, refine patient categorization for customized therapies, and provide novel insights into prospective vaccination targets for HBV-induced liver cancer. Also, our findings could help come up with new ways to treat HCC that take into account the different biological traits of younger and older patients. This would improve clinical outcomes and make HCC less of a problem around the world.

METHODOLOGY

Data collection

The raw gene expression data for this investigation were acquired from the Gene Expression Omnibus (GEO) database using specified keywords and selection criteria to extract significant expression profiles from tissue or clinical samples. GSE45267 was used for Hepatocellular Carcinoma (HCC). It has 87 samples of tumour and normal tissue that were paired together. 16 of the samples are from younger HCC (yHCC, ≤40 years) and 32 are from older HCC (eHCC, ≥40 years). This paired approach facilitated direct comparisons between neoplastic and normal tissues. Furthermore, the dataset GSE38941 was examined for Hepatitis B (HBV) expression data, which included samples from four HBV patients and ten healthy individuals to ascertain differentially expressed genes. Utilizing these datasets, transcriptome analyses were performed to identify age-related hub genes and molecular signatures in HBV-associated HCC, enhancing our understanding of prognostic biomarkers and prospective vaccine targets (Clough & Barrett, 2016).

Data pre-processing

The Series Matrix Files for GSE38941 and GSE45267 were acquired from the GEO database for further investigation. Before analysis, probe data from both datasets were transformed to standard gene symbols to synchronize gene identification with globally accepted nomenclature. The datasets were normalized using the Robust Multi-Array Average (RMA) technique in R software (version 26.0) to guarantee consistency and reduce any technical biases. This normalization procedure normalized gene expression data, guaranteeing consistency in size and distribution among the datasets (Agapito et al., 2022).

Identification of Differentially Expressed Genes (DEGs)

This work used GEO2R to discover differentially expressed genes (DEGs) in Hepatitis B vs. normal tissue samples and Young Hepatocellular Carcinoma (yHCC, ≤40) versus Elder Hepatocellular Carcinoma (eHCC, ≥40) samples. GEO2R produced a volcano plot, illustrating fold change in gene expression on the x-axis and statistical significance (p-value) on the y-axis. Differentially expressed genes (DEGs) were identified using a p-value threshold of <0.05 and an absolute log fold change above 2. Furthermore, we used FunRich V3.13 to illustrate the overlap and variability in DEGs across datasets, using a Venn diagram to emphasize common molecular signatures and pathways across the samples (Thirumani et al., 2024).

Protein-protein interaction and hub gene identification

Protein-protein interaction (PPI) analysis was performed using STRING by entering a list of proteins to visualize anticipated physical and functional correlations, with interactions exhibiting a total score over 0.08 deemed significant. The differentially expressed genes (DEGs) were used to create and display the protein-protein interaction (PPI) network using Cytoscape (version 3.5.1; http://www.cytoscape.org), with edges denoting protein linkages and edge widths reflecting interaction intensity. Hub genes were found using the Cytohubba plugin in Cytoscape, with nodes exhibiting a degree greater than 10 indicating their pivotal function in the network. This comprehensive method adeptly investigates intricate protein interactions and delineates essential biological processes (A. N. Hasan et al., 2015; Rao et al., 2014).

mRNA expression and survival analysis of hub genes

In silico techniques, including UALCAN, GEPIA, and KM plotter, were used to evaluate survival rates and gene expression relationships in young and elderly HCC patients. The Kaplan-Meier analysis, augmented by log-rank testing, substantiated the survival analysis. A statistically significant correlation was identified between gene expression levels and patient survival, meeting a significance criterion of $P < 0.05$. Patient data with gastric cancer, obtained from The Cancer Genome Atlas, were used for expression validation. The data, expressed as transcripts per million (TPM) values, facilitated the creation of two separate groups. To confirm the expression, data from hepatocellular carcinoma patients from The Cancer Genome Atlas were used, focusing on two groups: yHCC (young HCC) individuals and eHCC (older HCC) individuals. The data, represented as transcripts per million (TPM) values, enabled the identification of two distinct groups. The cohorts were shown with the GEPIA database. Patients with TPM readings below the upper quartile were classified into the low/medium expression group, while those over the upper quartile were designated as the high expression group.

Gene ontology and pathway enrichment analysis

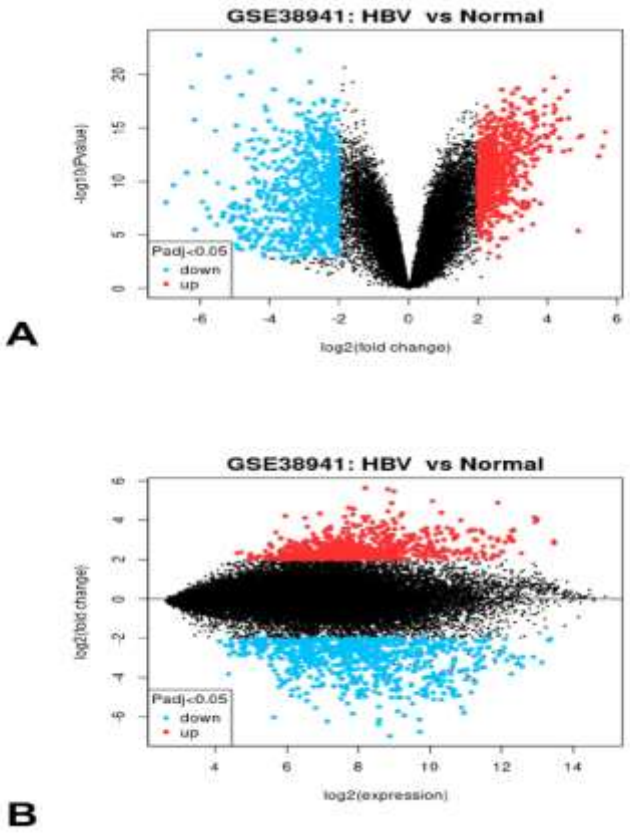
Gene Ontology (GO) and KEGG pathway enrichment analyses of differentially expressed genes (DEGs) were performed using the Database for Annotation, Visualization, and Integrated Discovery (DAVID, https://david.ncifcrf.gov/tools.jsp), with a significance level of $P < 0.05$. KEGG pathway analysis revealed highly enriched pathways linked to DEGs, whereas pathway crosstalk analysis used a Benjamini-Hochberg adjusted P-value < 0.05, along with both a Jaccard coefficient and an overlap coefficient > 0.5 as statistical criteria. This thorough approach emphasizes the role of DEGs in essential biological processes and regulatory networks.

RESULT

Identification of Differentially Expressed Genes (DEGs) in Hepatitis B and Hepatocellular Carcinoma (HCC) Across Age Groups

In the analysis of the Hepatitis B dataset GSE38941, a total of 1322 differentially expressed genes (DEGs) were identified via GEO2R using the Limma package. The selection criteria included an adjusted p-value of less than 0.05 and a log fold change greater than 1, resulting in the creation of volcano plots (Figures 1A and 1B). Similarly, the examination of the HCC dataset GSE45267 identified 550 differentially expressed genes in elder HCC (eHCC) and

282 differentially expressed genes in younger HCC (yHCC) samples (Figures 2A and 2B). Using FunRich V3.13 software, we identified 249 overlapping DEGs between eHCC and yHCC (Figure 2C). Furthermore, the most regulated genes in eHCC and yHCC are illustrated in (Figures 3A and 3B), respectively.



Figures 1A and 1B

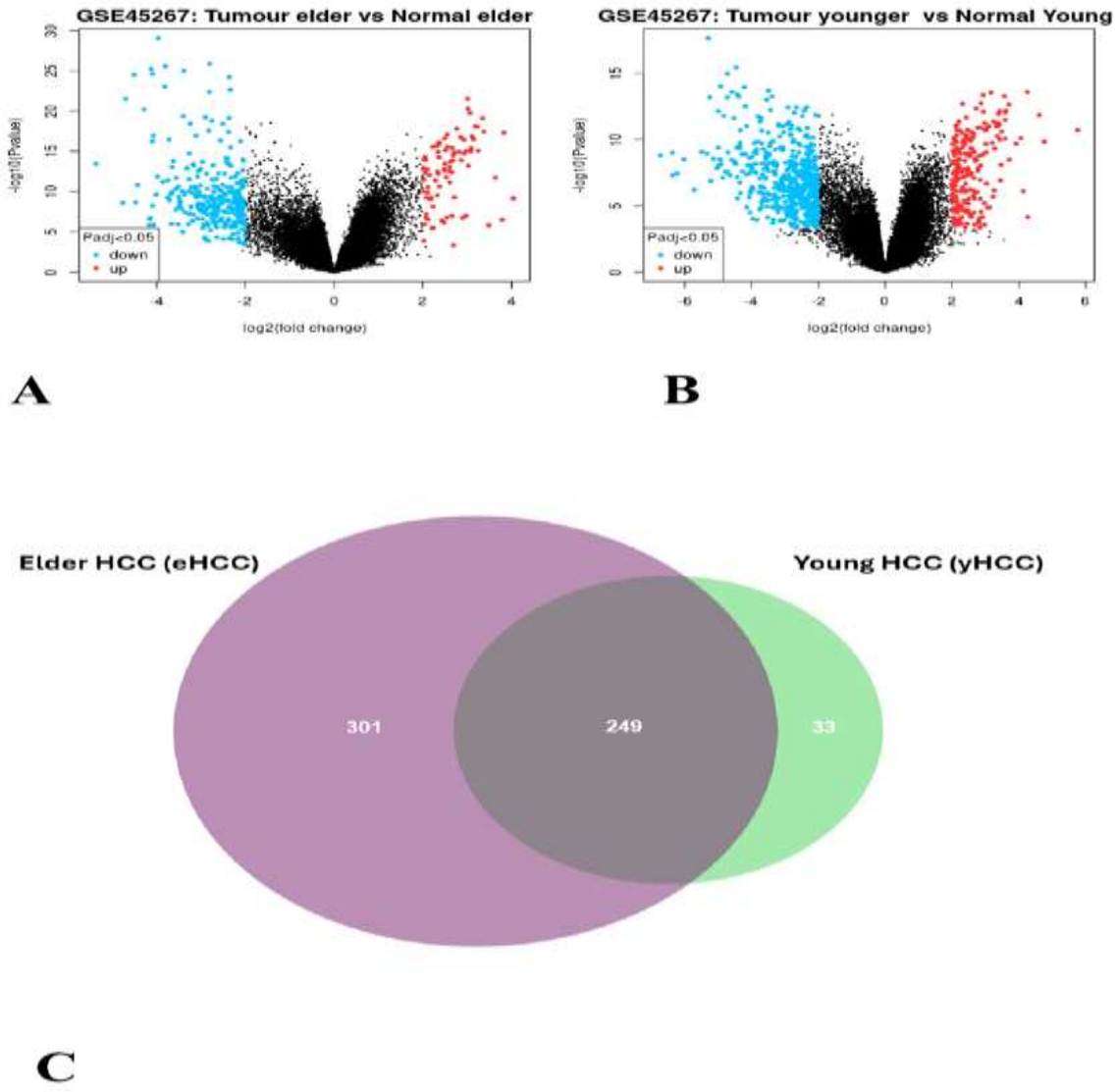
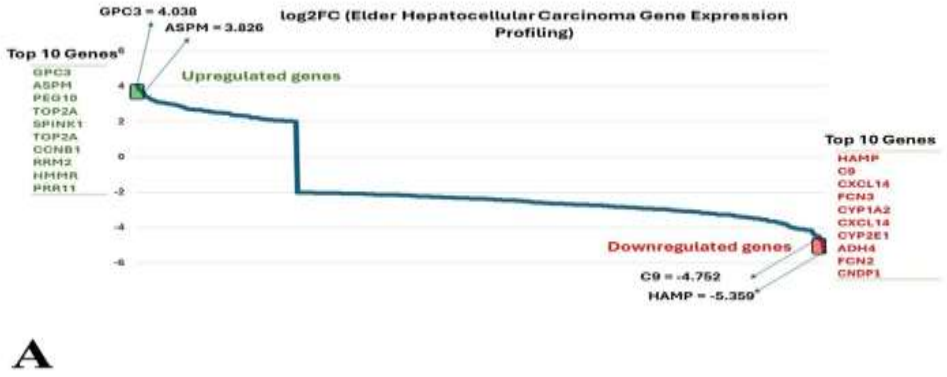
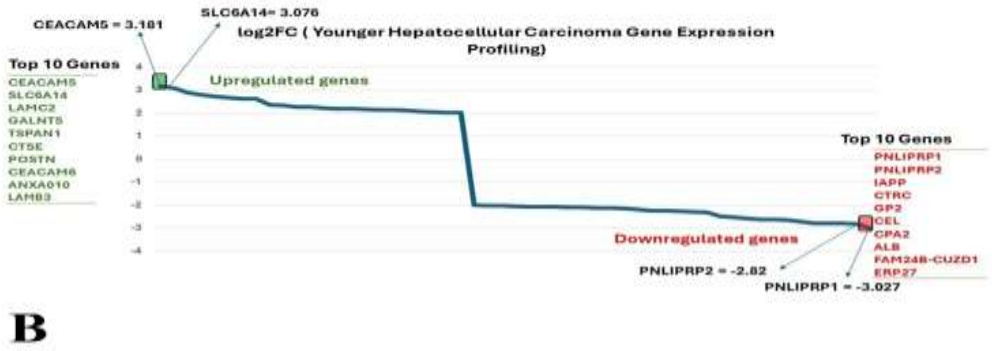


Figure 2A, 2B and 2C



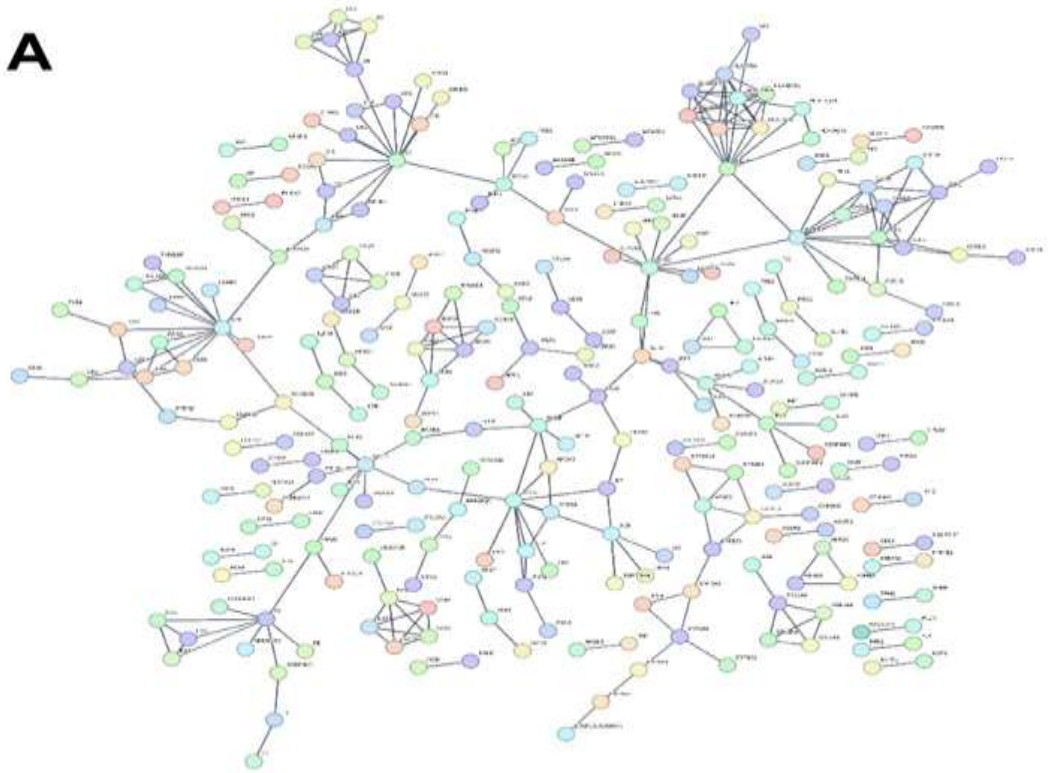


Figures 3A and 3B

PPI network construction and hub gene identification

A network of protein-protein interactions was established for differentially expressed genes through STRING analysis and subsequently visualized using Cytoscape. In (Figure 4A), a total of 1322 differentially expressed genes (DEGs) led to the identification of 914 interconnected proteins. Furthermore, four hub genes—SYK, C3, CD44, and IGLL5 were determined through the application of Bottleneck, closeness, and degree centrality methods (Figure 4B and 4C). In the same way, 249 DEGs that overlapped helped find 227 proteins that were linked to each other (Figure 5A). Eight hub genes (CDK1, UBE2C, BUB1B, CCNB1, CDC20, CCNA2, MAD2L1, and BIRC5) were found using MCC, closeness, and degree centrality methods (Figure 5B and 5C). All eight hub genes were notably upregulated, suggesting their potential importance in the development of HCC.

Figure 4A



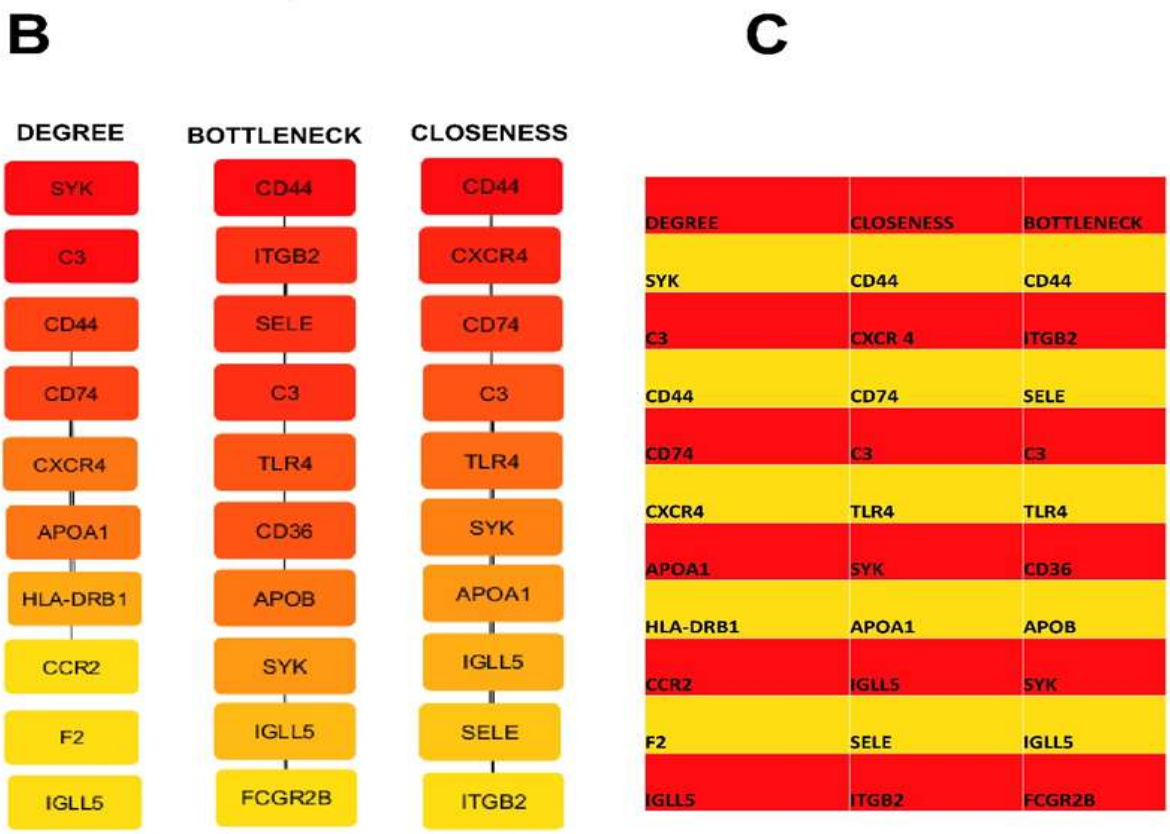
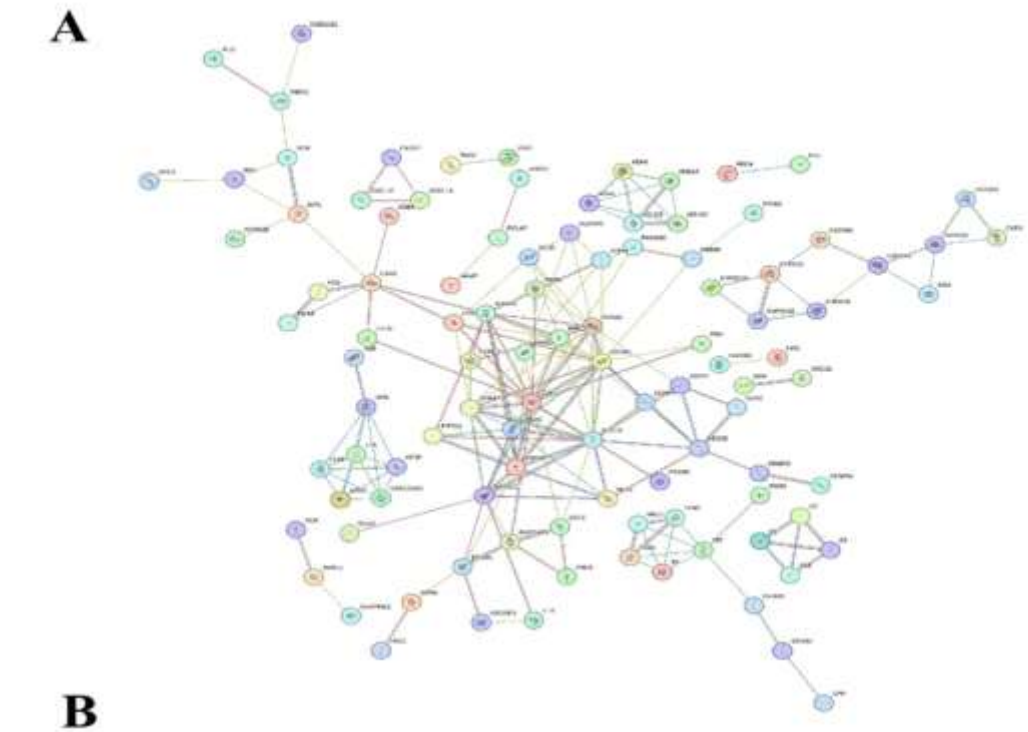
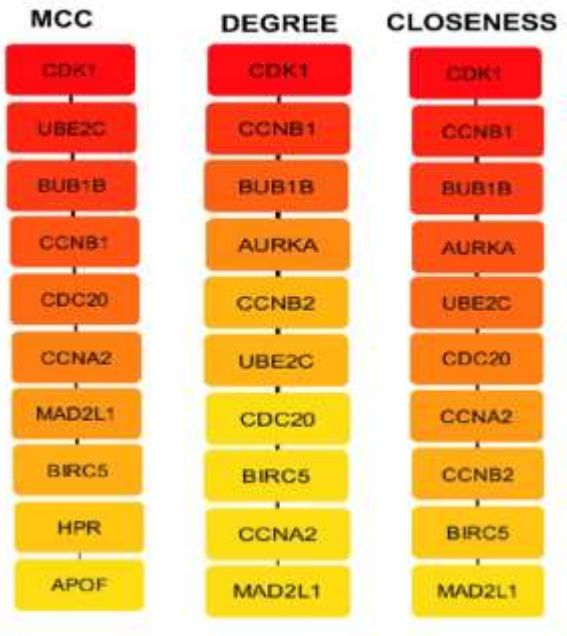


Figure 4B and 4C



B



C

MCC	DEGREE	CLOSENESS
CDK1	CDK1	CDK1
UBE2C	CCNB1	CCNB1
BUB1B	BUB1B	BUB1B
CCNB1	AURKA	AURKA
CDC20	CCNB2	UBE2C
CCNA2	UBE2C	CDC20
MAD2L1	CDC20	CCNA2
BIRC5	BIRC5	CCNB2
HPR	CCNA2	BIRC5
APOF	MAD2L1	MAD2L1

Figure 5A, 5B and 5C

Gene ontology and KEGG pathway analysis of DEGs

The examination of differentially expressed genes (DEGs) using Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways uncovers complex biological processes and pathways involved in the pathogenesis of hepatocellular carcinoma (HCC) and gastric adenocarcinoma (GAC). The Gene Ontology analysis classifies differentially expressed genes into three primary categories: biological process, cellular component, and molecular function. Both hepatocellular carcinoma (HCC) and gastric adenocarcinoma (GAC) demonstrate considerable modifications in metabolic pathways, encompassing carboxylic acid and organic acid biosynthesis and catabolism, steroid and fatty acid metabolism, and amino acid metabolism, all of which are essential for cellular energy production, biosynthesis, and tumour proliferation (Figures 6A & 9A). The dysregulation of lipid metabolism, highlighted by fatty acid and steroid metabolic pathways, indicates the tumour's need for lipids for survival, while amino acid metabolism facilitates tumour growth and protein production. Furthermore, humoral immune response pathways are crucial, especially in GAC, affecting antibody synthesis and immunological control.

At the cellular level, critical structures like blood microparticles, extracellular matrix components, and chromosomal regions play a vital role in tumour growth. The existence of blood microparticles and extracellular matrix components in both malignancies indicates significant interactions between tumours and their microenvironments that facilitate invasion and metastasis. In HCC, certain chromosomal areas and kinetochores underscore genetic instability, a characteristic of malignancy (Figure 6B). In GAC, elements like the collagen-rich extracellular matrix, vesicle lumen, secretory granule lumen, and plasma lipoprotein particles are associated with inflammation, fibrosis, and viral protein trafficking, especially regarding hepatitis B infection (Figure 9B). The endoplasmic reticulum lumen is essential for protein folding and processing, while the extracellular aspect of the plasma membrane is a critical locus for viral entrance and immunological interactions. At the molecular function level, oxidoreductase activity, heme binding, monooxygenase activity, sulfur compound binding, and lipid transporter activity are crucial for the control of oxidative stress, metabolic detoxification, and immunological responses (Figures 6C & 9C). The detection of oxidoreductase and monooxygenase activities highlights the significance of redox signalling and metabolic detoxification, while lipid transporter activity is essential for sustaining lipid homeostasis, with its dysregulation leading to hepatic steatosis. Furthermore, the binding activities of steroids and enzyme inhibitors underscore the impact of these genes on hormonal control and viral propagation. KEGG pathway analysis elucidates critical pathways implicated in disease development. In hepatocellular carcinoma (HCC), pathways including complement cascades, retinol metabolism, drug metabolism via cytochrome P450, xenobiotic metabolism, chemical carcinogenesis-DNA adducts, cell cycle regulation, tryptophan metabolism, steroid hormone biosynthesis, bile secretion, and p53 signalling highlight the intricate interplay of metabolic and regulatory mechanisms in hepatocarcinogenesis (Figures 7 & 8). In GAC, pathways such as complement and coagulation cascades, cholesterol and bile acid metabolism, cytochrome P450, hematopoietic cell lineage, rheumatoid arthritis, leishmaniasis, and Staphylococcus aureus infection underscore immune system dysregulation and metabolic disturbances associated with chronic hepatitis B infection (Figures 10 & 11). These results together elucidate the molecular heterogeneity of HCC and GAC, revealing prospective treatment targets and biomarkers for early diagnosis and precision medicine. This work integrates GO and KEGG analyses to elucidate critical genetic and molecular pathways involved in hepatocellular and gastric carcinogenesis, providing significant insights for future research and intervention.

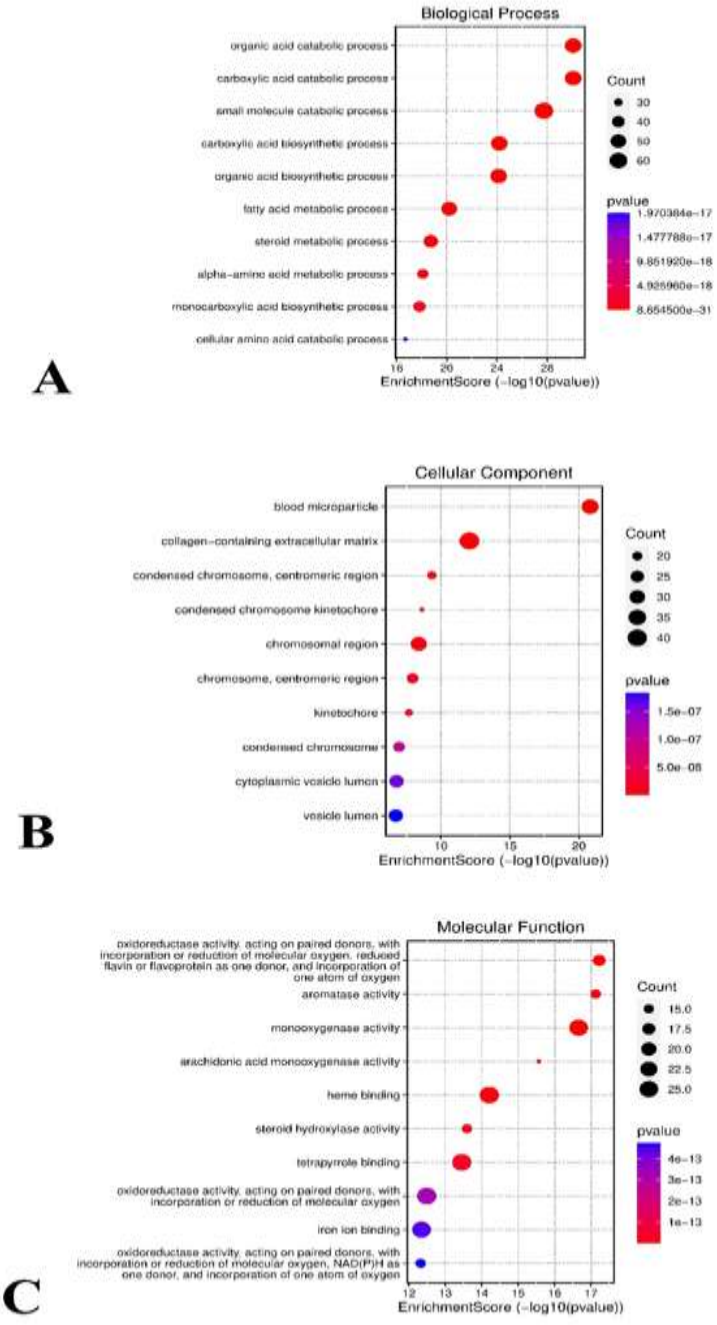


Figure 6A, 6B and 6C

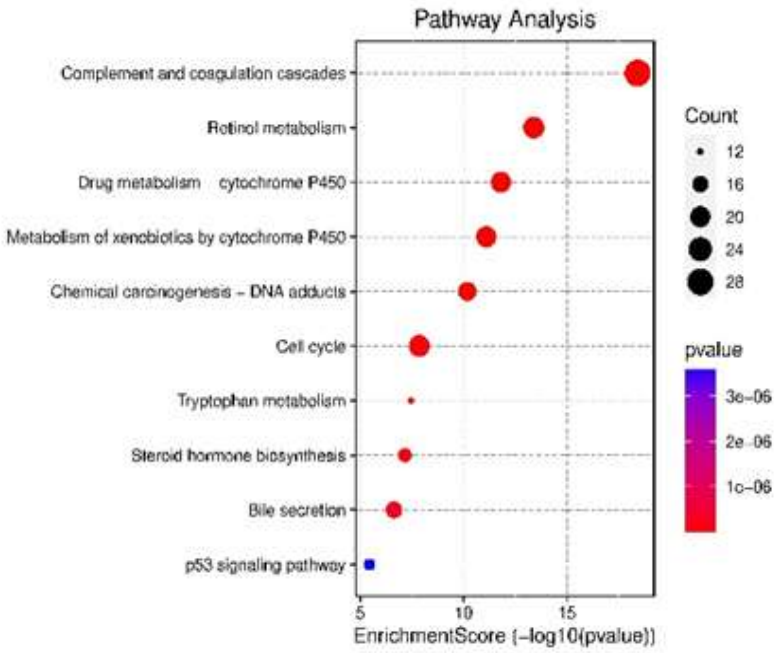

Pathway Analysis
Complement and coagulation cascades
Retinol metabolism
Drug metabolism - cytochrome P450
Metabolism of xenobiotics by cytochrome P450
Chemical carcinogenesis - DNA adducts
Cell cycle
Tryptophan metabolism
Steroid hormone biosynthesis
Bile secretion
p53 signaling pathway
5
10
15
EnrichmentScore (-log10(pvalue))
Count
12
16
20
24
28
pvalue
3e-06
2e-06
1e-06

Figure 7

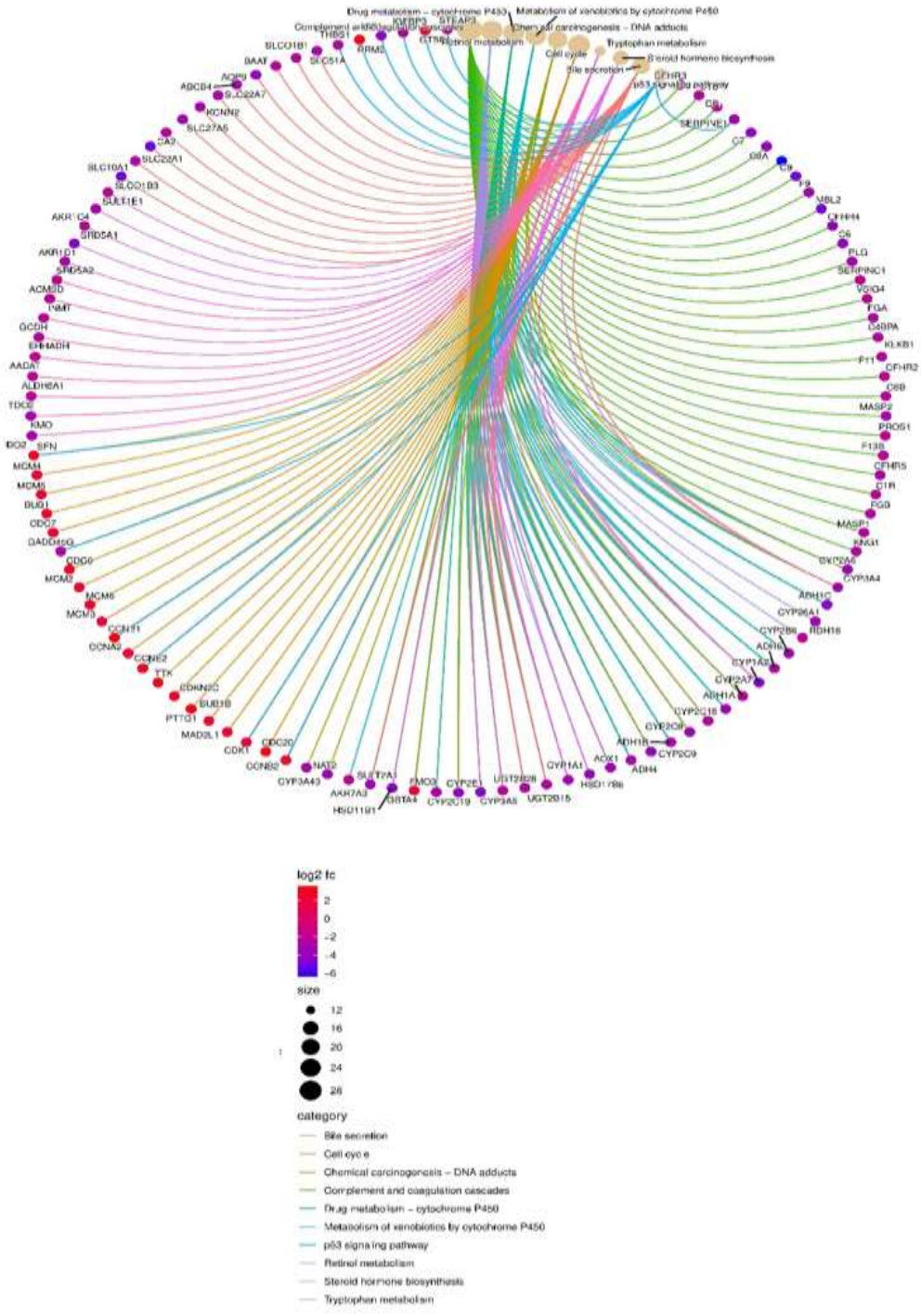

log2 fc
2
0
-2
-4
-6
size
12
16
20
24
28
category
Bile secretion
Cell cycle
Chemical carcinogenesis – DNA adducts
Complement and coagulation cascades
Drug metabolism – cytochrome P450
Metabolism of xenobiotics by cytochrome P450
p53 signaling pathway
Retinol metabolism
Steroid hormone biosynthesis
Tryptophan metabolism

Figure 8

Figure 9A, 9B and 9C

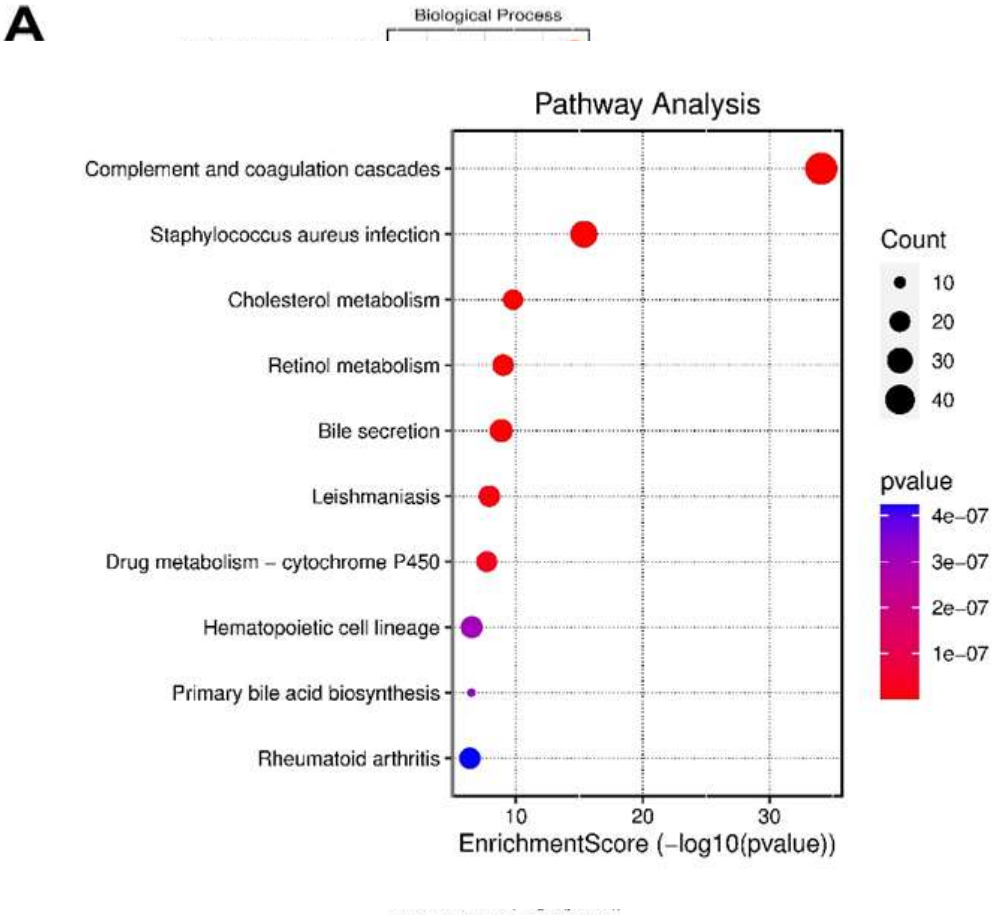
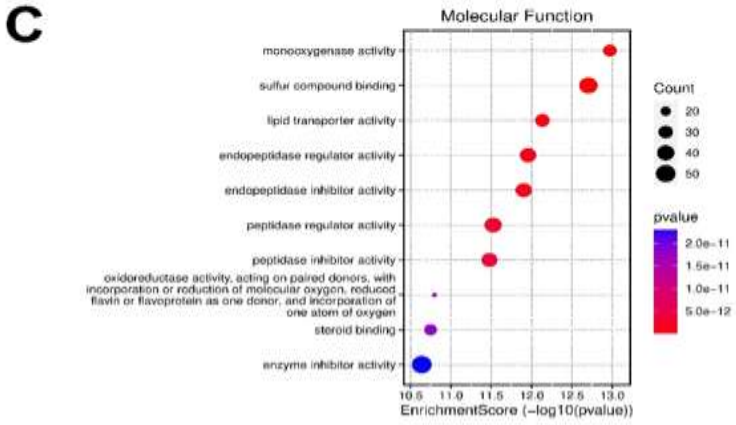


Figure 10



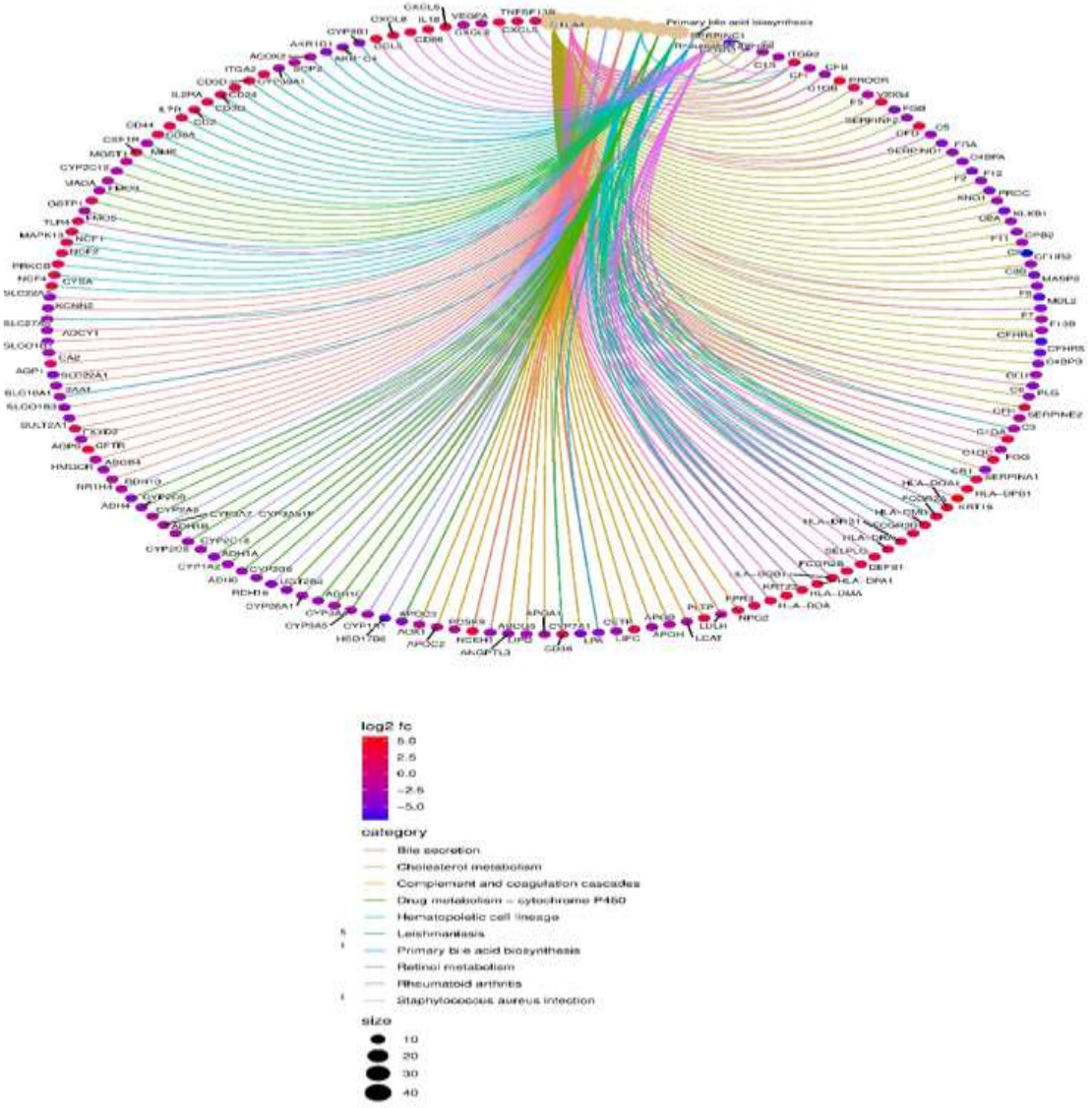
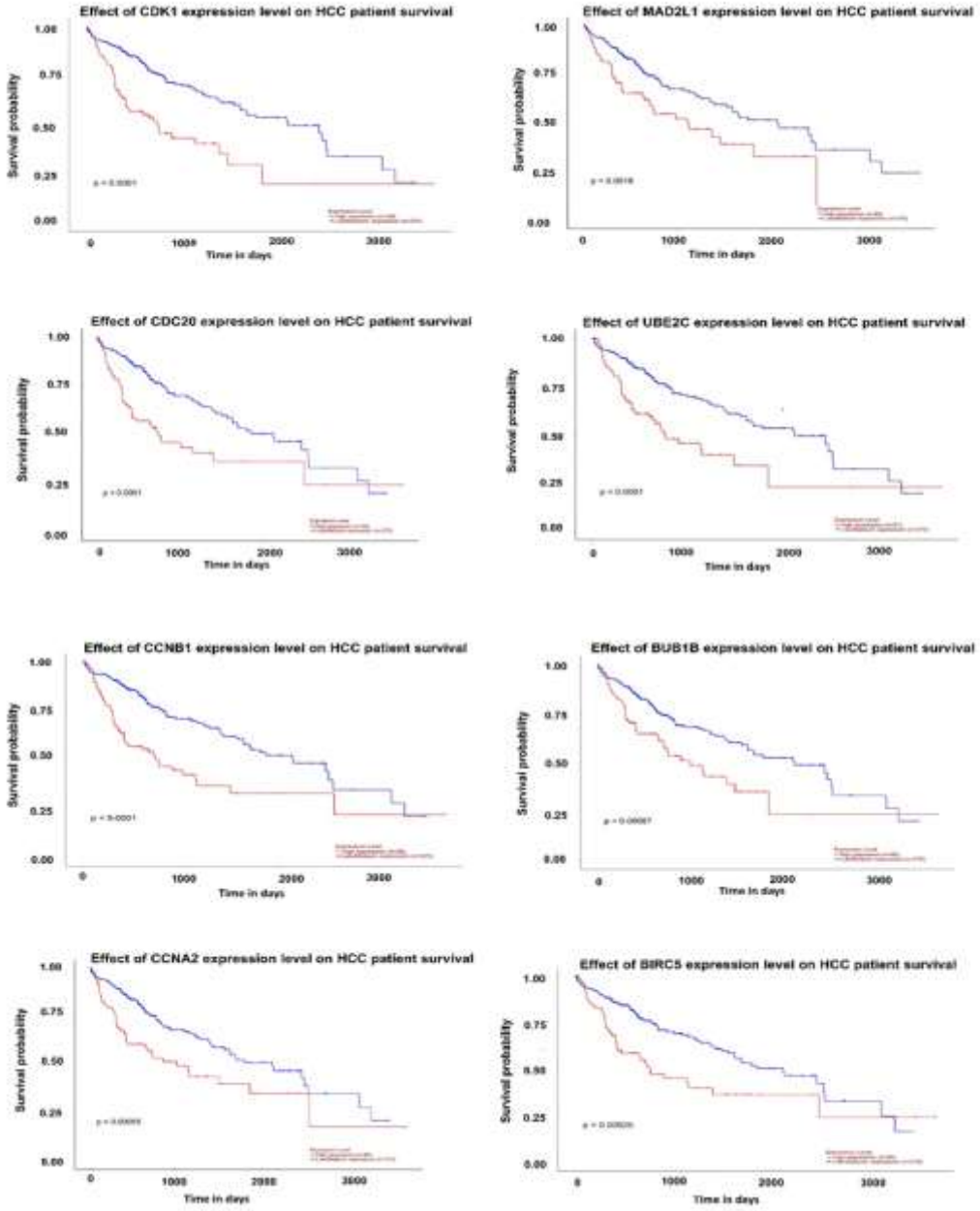


Figure 11

Survival analysis of hub genes

In the study of hepatocellular carcinoma, assessing the expression levels and conducting survival analysis of hub genes is essential for comprehending the disease's molecular dynamics. The survival study demonstrated substantial relationships between gene expression levels and patient survival outcomes, as evidenced by the p-value significance for eight hub genes in HCC common to young and elderly patients. Genes include BIRC5 ($p = 0.00025$), BUB1B ($p = 0.00087$), CCNA2 ($p = 0.00055$), MAD2L1 ($p = 0.0018$), and CDC20, CCNB1, CDK1, and UBE2C ($p < 0.0001$) are shown (Figure 12). The hub genes had extremely significant p-values, indicating their potential as prognostic indicators in HCC. The p-value indicates that the dysregulation of these hub genes, which are implicated in cell cycle



regulation (BUB1B, CDC20, CCNB1, CDK1) and mitotic checkpoint control (MAD2L1), signifies disturbances in the cell division process that lead to tumor development and metastasis. Additionally, genes implicated in apoptosis control, such as BIRC5, and those involved in ubiquitin-mediated protein degradation, including UBE2C, may affect tumor cell viability and treatment efficacy. The finding of these hub genes as prognostic indicators underscores their potential therapeutic significance in forecasting patient outcomes and informing tailored treatment strategies for HCC.

Figure 12

Verification of mRNA expression analysis

Employing the GEPIA web server, a technology esteemed for cancer data processing, our study focused on eight critical hub genes. The boxplot demonstrated a consistent elevation of the eight hub genes: CDK1, BIRC5, BUB1B, CCNA2, CCNB1, CDC20, MAD2L1, and UBE2C in hepatocellular carcinoma (HCC) (Figure 13). The increase indicates their possible role in facilitating tumor development and progression. CDK1, which regulates the cell cycle, and BIRC5, associated with apoptosis suppression, may facilitate unchecked cell proliferation and survival. Likewise, the upregulation of BUB1B, CDC20, CCNB1, and MAD2L1, which are linked to mitotic checkpoint regulation, may promote abnormal cell division and genomic instability. Furthermore, elevated expression of UBE2C, which is implicated in ubiquitin-mediated protein degradation, may amplify carcinogenic signalling pathways. These results underscore the importance of hub genes in HCC etiology and their potential as therapeutic targets. Our research using the UALCAN web server, a cancer data analysis platform, identified the expression patterns of eight hub genes across several cancer stages, revealing significant tendencies in HCC development (Figure 14). MAD2L1, CCNA2, BIRC5, and BUB1B demonstrate elevated expression only in stage three of cancer, indicating their involvement in facilitating advanced tumour proliferation and metastasis. In contrast, CDK1, CDC20, CCNB1, and UBE2C exhibit increased expression in both stages two and three of cancer, suggesting their role in the initial phases of carcinogenesis and the advancement to later stages. This data indicates a dynamic modulation of hub gene expression throughout HCC development, with some genes facilitating tumour start and early proliferation, while others gain prominence as the malignancy progresses to advanced stages. Comprehending the stage-specific expression patterns of these hub genes elucidates the molecular processes driving HCC growth and progression, aiding in the discovery of possible prognostic indicators and treatment targets.

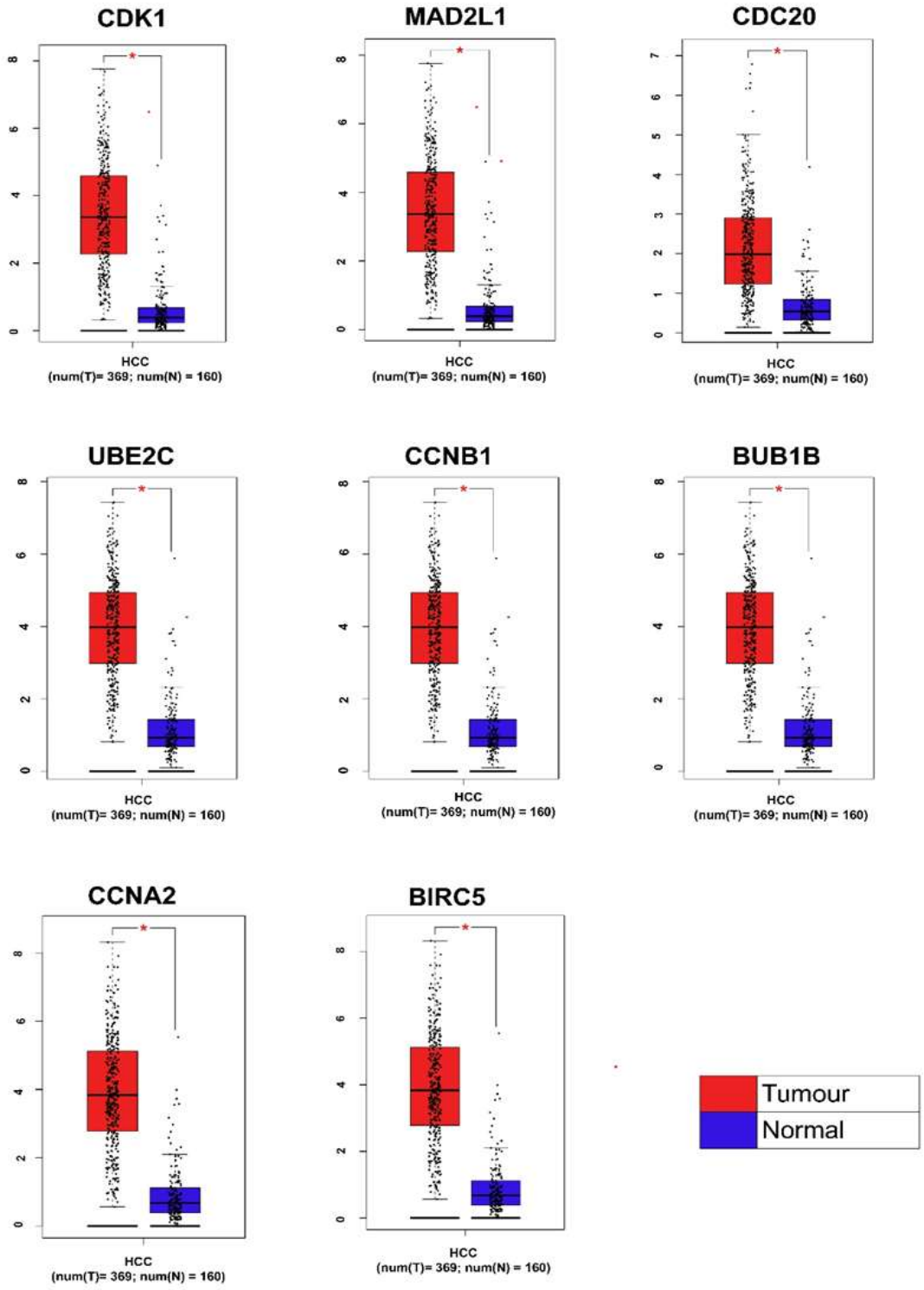

CDK1
MAD2L1
CDC20
UBE2C
CCNB1
BUB1B
CCNA2
BIRC5
HCC
(num(T)= 369; num(N) = 160)
Tumour
Normal

Figure 13

Figure 14

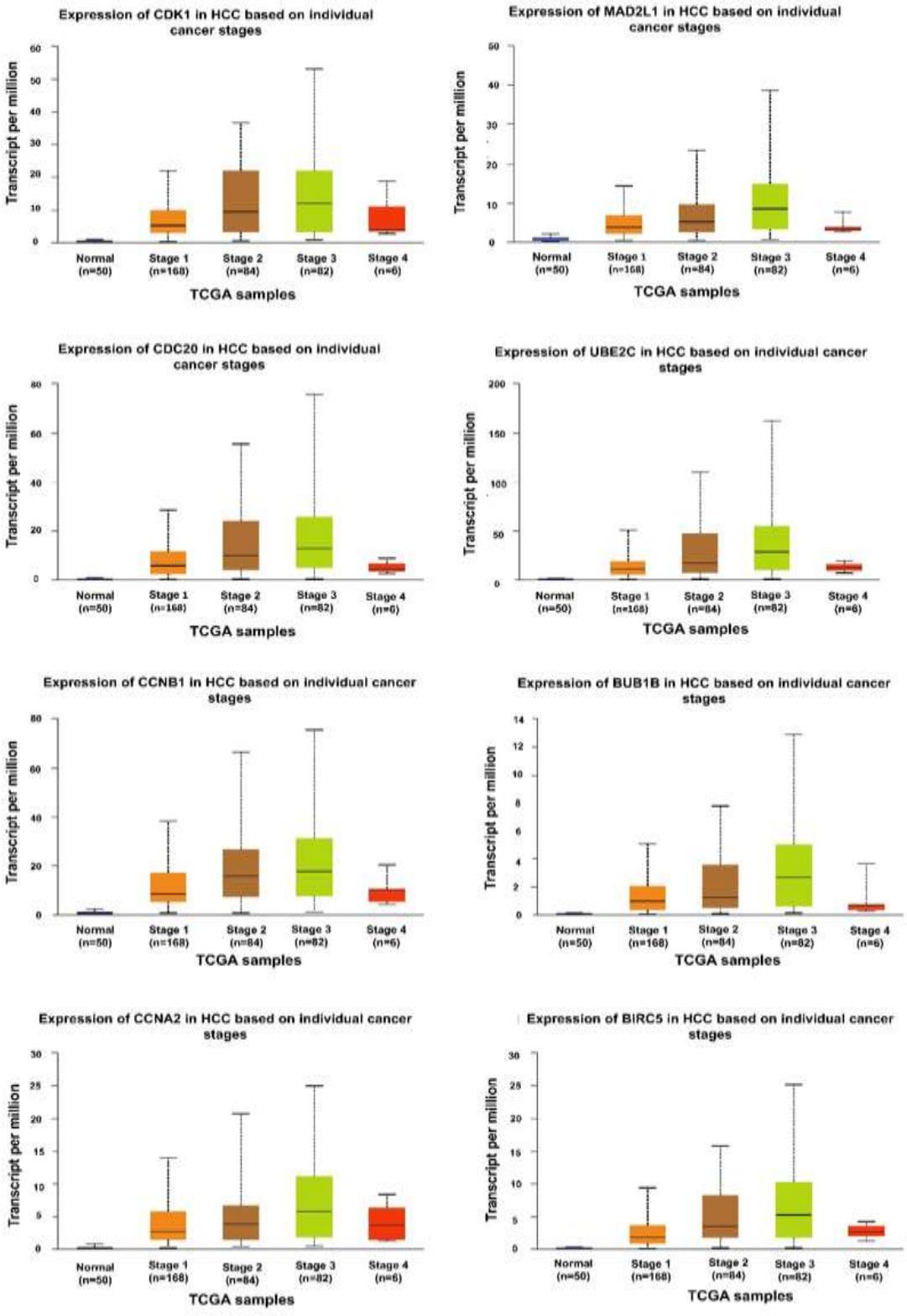

Expression of CDK1 in HCC based on individual cancer stages
Expression of MAD2L1 in HCC based on individual cancer stages
Expression of CDC20 in HCC based on individual cancer stages
Expression of UBE2C in HCC based on individual cancer stages
Expression of CCNB1 in HCC based on individual cancer stages
Expression of BUB1B in HCC based on individual cancer stages
Expression of CCNA2 in HCC based on individual cancer stages
Expression of BIRC5 in HCC based on individual cancer stages
Transcript per million
Normal (n=50)
Stage 1 (n=168)
Stage 2 (n=84)
Stage 3 (n=82)
Stage 4 (n=6)
TCGA samples

Figure 15 The

GENE SYMBOL	GENE TITLE	LOG 2FC	P VALUE	CHROMOSOME NUMBER
GPC3	glypican 3	4.038	9.17	chrX:133,535,745-133,987,100
ASPM	abnormal spindle microtubule assembly	3.826	17.326	chr1:197,084,121-197,146,694
PEG10	paternally expressed 10	3.79	6.506	chr7:94,656,325-94,669,695
TOP2A	topoisomerase (DNA) II alpha	3.633	11.72	chr17:40,388,525-40,417,896
SPINK1	serine peptidase inhibitor, Kazal type 1	3.487	5.814	chr5:147,824,572-147,839,196
CCNB1	cyclin B1	3.348	19.1	chr5:69,167,135-69,178,245
RRM2	ribonucleotide reductase regulatory subunit M2	3.246	15.081	chr2:10,120,698-10,211,725
HMMR	hyaluronan mediated motility receptor	3.227	18.06	chr5:163,460,203-163,491,941
PRR11	proline rich 11	3.184	18.117	chr17:59,155,732-59,206,709
CDKN3	cyclin dependent kinase inhibitor 3	3.145	16.517	chr14:54,396,849-54,420,218
TTK	TTK protein kinase	3.11	16.881	chr6:80,003,887-80,042,527
ANLN	anillin actin binding protein	3.088	15.388	chr7:36,389,821-36,453,791
RACGAP1	Rac GTPase activating protein 1	3.061	19.772	chr12:49,976,923-50,033,440
DTL	denticleless E3 ubiquitin protein ligase homolog	3.027	13.209	chr1:212,035,553-212,107,400
ECT2	epithelial cell transforming 2	3.016	20.313	chr3:172,750,682-172,829,265
PRC1	protein regulator of cytokinesis 1	3.009	21.531	chr15:90,966,040-90,995,629
GINS1	GINS complex subunit 1	2.977	15.069	chr20:25,391,008-25,452,700
SULT1C2	sulfotransferase family 1C member 2	2.97	7.01	chr2:108,288,639-108,309,915
AURKA	aurora kinase A	2.931	14.736	chr20:56,369,389-56,392,337
NDC80	NDC80, kinetochore complex component	2.925	16.426	chr18:2,571,557-2,616,635

depiction of Differentially Expressed Genes of top 20 upregulated genes and their chromosome number correlation for targeted therapy

Figure 16 The depiction of the top 20 upregulated Chromosomal Localization of Differentially Expressed Genes in Younger Hepatocellular Carcinoma (HCC) Patients

CHROMOSOME NUMBER FOR YOUNGER HCC

GENE SYMBOL	GENE TITLE	LOG2FC	P VALUE	CHROMOSOME NUMBER
CEACAM5	carcinoembryonic antigen related cell adhesion molecule 5	3.181	12.665	chr19:41,708,585-41,730,433
SLC6A14	solute carrier family 6 member 14	3.076	14.577	chrX:116,436,606-116,461,458
LAMC2	laminin subunit gamma 2	2.902	18.153	chr1:183,186,238-183,258,968
GALNT5	polypeptide N-acetylgalactosaminyltransferase 5	2.789	15.442	chr2:157,257,705-157,318,491
TSPAN1	tetraspanin 1	2.727	17.604	chr1:46,175,073-46,196,489
CTSE	cathepsin E	2.684	14.222	chr1:206,008,535-206,023,909
POSTN	periostin	2.63	12.17	chr13:37,562,583-37,598,844
CEACAM6	carcinoembryonic antigen related cell adhesion molecule 6	2.627	11.974	chr19:41,750,977-41,772,211
ANXA10	annexin A10	2.358	10.282	chr4:168,092,537-168,187,736
LAMB3	laminin subunit beta 3	2.344	17.925	chr1:209,614,870-209,652,425
ITGA2	integrin subunit alpha 2	2.266	13.249	chr5:52,989,340-53,094,779
TMPRSS4	transmembrane protease, serine 4	2.257	15.545	chr11:118,077,012-118,125,505
FN1	fibronectin 1	2.211	11.488	chr2:215,360,440-215,436,073
COL11A1	collagen type XI alpha 1 chain	2.195	12.516	chr1:102,876,467-103,108,872
SERPINB5	serpin family B member 5	2.183	13.901	chr18:63,476,958-63,505,085
DPCR1	diffuse panbronchiolitis critical region 1	2.158	8.978	chr6:30,934,523-30,954,221
AGR2	anterior gradient 2, protein disulphide isomerase family member	2.137	12.245	chr7:16,791,811-16,833,433

DISCUSSION

The recent finding of age-related hub genes in Hepatitis B-associated hepatocellular carcinoma (HBV-HCC) offers essential insights into the molecular processes underlying disease development and possible prognostic indicators. This work used transcriptome analysis to investigate differentially expressed genes (DEGs) in HBV-associated HCC across age demographics, elucidating the variability of gene expression in young (yHCC) and elderly (eHCC) patients (Kalaki et al., 2024; J. Zhang et al., 2021). The findings demonstrated unique gene expression patterns, suggesting that tumor biology varies considerably with age, potentially affecting prognosis, therapy, and vaccine development. Our investigation revealed 1,322 differentially expressed genes (DEGs) in HBV samples relative to normal tissues, 550 DEGs in early hepatocellular carcinoma (eHCC), and 282 DEGs in young hepatocellular carcinoma (yHCC) (Xie et al., 2020). The intersection of 249 differentially expressed genes (DEGs) between early hepatocellular carcinoma (eHCC) and young hepatocellular carcinoma (yHCC) underscored common molecular characteristics. Utilizing STRING and Cytoscape, key hub genes (CDK1, UBE2C, BUB1B, CCNB1, CDC20, CCNA2, MAD2L1, and BIRC5) were discerned, which were markedly elevated in HCC and contributed considerably to tumor growth, cell cycle dysregulation, and immune modulation (M. A. M. Hasan et al., 2023). The Gene Ontology (GO) and KEGG pathway enrichment analyses yielded further functional insights, indicating that these genes participate in metabolic pathways, immune response regulation, and cell cycle control, which are critical activities implicated in HCC pathogenesis. The survival study highlighted the predictive importance of these hub genes, demonstrating their correlation with patient outcomes. Genes including BIRC5, CCNA2, and BUB1B were notably significant, since their elevated expression was associated with reduced survival rates (Hamdy et al., 2023; Y. Shen et al., 2024). The validation with GEPIA and UALCAN corroborated the differential expression of these genes at various phases of HCC, hence reinforcing their potential as biomarkers for disease progression and therapeutic targets. These results correspond with other research highlighting the significance of cell cycle regulators and immune-modulating genes in HCC. Research has emphasized the overexpression of BIRC5 (survivin) in hepatocellular carcinoma (HCC) and its correlation with the suppression of apoptosis and chemoresistance (Jiang et al., 2021). CDK1 and CCNA2 have been associated with unregulated cell proliferation and tumor advancement in liver malignancies. Age-related variations in gene expression patterns correspond with findings indicating that younger HCC patients have more aggressive tumor phenotypes, while older patients show immunosuppressive tumor microenvironments attributable to age-related changes in immune surveillance. The implications of these discoveries go beyond biomarker identification to possible medicinal uses. The discovery of age-related hub genes offers prospects for targeted therapy regimens customized for distinct age groups, hence improving precision medicine techniques in HCC care. Moreover, the findings from the research may enhance vaccine development efforts by pinpointing molecular targets essential for immune regulation in HBV-associated liver cancer. Notwithstanding these merits, the research has drawbacks. The datasets used were sourced from publicly accessible materials, potentially introducing selection biases. Furthermore, experimental confirmation of the discovered hub genes in clinical specimens is essential to ascertain their biological significance. Subsequent study must concentrate on functional investigations to clarify the specific molecular processes of these genes in hepatocarcinogenesis and their potential as therapeutic targets. Several previous studies have explored gene expression profiling in breast cancer to identify molecular markers and potential therapeutic targets. (Blows et al., 2010) initiated the categorization of breast cancer subtypes according to gene expression profiles, resulting in the recognition of luminal, basal-like, and HER2-enriched subtypes. Their research established a basis for comprehending the molecular heterogeneity of breast cancer. (Asleh et al., 2022) advanced this research by demonstrating that intrinsic gene expression patterns might forecast patient survival outcomes and therapeutic responses. Their results underscored the therapeutic significance of transcriptome profiling in the treatment of breast cancer . Recently, The Cancer Genome Atlas (TCGA) research group has supplied comprehensive genomic datasets that enable broad comparisons between normal and malignant breast tissues. Their research has uncovered new genetic abnormalities, epigenetic alterations, and disrupted signaling pathways linked to tumor growth (Tomczak et al., 2015). Furthermore, studies by (Ciriello et al., 2015) have helped us learn more about the molecular subtypes and genetic changes that make breast cancer so different. This research has highlighted the significance of combining transcriptome data with clinical characteristics to enhance prognostic models. Our results validate previous research and provide additional insights by identifying prospective candidate biomarkers unique to our dataset. There needs to be more testing with different groups of patients and functional assays to make sure that these biomarkers are useful for diagnosing and treating breast cancer.

CONCLUSION

This research effectively identified age-related hub genes in Hepatitis B-associated hepatocellular carcinoma (HBV-HCC) by transcriptome analysis. Utilizing differential gene expression analysis, protein-protein interaction networks, and survival analysis, we found pivotal hub genes, such as CDK1, BIRC5, BUB1B, CCNA2, CCNB1, CDC20, MAD2L1, and UBE2C, which are integral to hepatocarcinogenesis. These genes are mostly implicated in cell cycle regulation, mitotic checkpoint control, and apoptosis inhibition, suggesting their potential as prognostic biomarkers and therapeutic targets. The research underscores divergent gene expression profiles between younger (≤40 years) and older (≥40 years) HCC patients, highlighting the influence of age on tumor biology. The overexpression of hub genes across several cancer stages indicates their role in both early and advanced tumor growth, potentially facilitating the development of stage-specific diagnostic and therapeutic approaches. Gene ontology and KEGG pathway studies identified critical metabolic and immunological-related pathways that facilitate tumor formation, underscoring the significance of lipid metabolism, oxidative stress management, and immune modulation in HCC progression. Our results provide significant insights into the molecular heterogeneity of HBV-HCC, facilitating individualized treatment strategies. The found hub genes may function as prospective targets for early diagnosis and therapeutic intervention, enhancing patient stratification and clinical outcomes. Furthermore, comprehending age-related molecular alterations in HCC may aid in the formulation of age-specific prognostic markers and vaccination tactics, possibly alleviating the worldwide impact of HBV-related liver cancer. Subsequent research should concentrate on verifying these hub genes in more extensive patient cohorts and investigating their functional involvement in hepatocarcinogenesis. Examining the relationship among age, immunological response, and HBV infection in the course of HCC may provide innovative therapeutic developments, hence improving precision medicine strategies for liver cancer treatment.

REFERENCE

1. Abdelhamed, W., & El-Kassas, M. (2024). Hepatitis B virus as a risk factor for hepatocellular carcinoma: There is still much work to do. Liver Research, 8(2), 83–90. https://doi.org/10.1016/j.livres.2024.05.004
2. Agapito, G., Milano, M., & Cannataro, M. (2022). A statistical network pre-processing method to improve relevance and significance of gene lists in microarray gene expression studies. BMC Bioinformatics, 23(Suppl 6), 393. https://doi.org/10.1186/s12859-022-04936-z
3. Asafo-Agyei, K. O., & Samant, H. (2025). Hepatocellular carcinoma. In StatPearls. StatPearls Publishing. https://www.ncbi.nlm.nih.gov/books/NBK559177/
4. Asleh, K., Riaz, N., & Nielsen, T. O. (2022). Heterogeneity of triple negative breast cancer: Current advances in subtyping and treatment implications. Journal of Experimental & Clinical Cancer Research: CR, 41(1), 265. https://doi.org/10.1186/s13046-022-02476-1
5. Atyah, M., Yin, Y.-R., Zhou, C.-H., Zhou, Q., Chen, W.-Y., Dong, Q.-Z., & Ren, N. (2018). Integrated analysis of the impact of age on genetic and clinical aspects of hepatocellular carcinoma. Aging, 10(8), 2079–2097. https://doi.org/10.18632/aging.101531
6. Blows, F. M., Driver, K. E., Schmidt, M. K., Broeks, A., van Leeuwen, F. E., Wesseling, J., Cheang, M. C., Gelmon, K., Nielsen, T. O., Blomqvist, C., Heikkilä, P., Heikkinen, T., Nevanlinna, H., Akslen, L. A., Bégin, L. R., Foulkes, W. D., Couch, F. J., Wang, X., Cafourek, V., … Huntsman, D. (2010). Subtyping of breast cancer by immunohistochemistry to investigate a relationship between subtype and short and long term survival: a collaborative analysis of data for 10,159 cases from 12 studies. PLoS Medicine, 7(5), e1000279. https://doi.org/10.1371/journal.pmed.1000279
7. Chan, Y.-T., Zhang, C., Wu, J., Lu, P., Xu, L., Yuan, H., Feng, Y., Chen, Z.-S., & Wang, N. (2024). Biomarkers for diagnosis and therapeutic options in hepatocellular carcinoma. Molecular Cancer, 23(1), 189. https://doi.org/10.1186/s12943-024-02101-z
8. Ciriello, G., Gatza, M. L., Beck, A. H., Wilkerson, M. D., Rhie, S. K., Pastore, A., Zhang, H., McLellan, M., Yau, C., Kandoth, C., Bowlby, R., Shen, H., Hayat, S., Fieldhouse, R., Lester, S. C., Tse, G. M. K., Factor, R. E., Collins, L. C., Allison, K. H., … Perou, C. M. (2015). Comprehensive molecular portraits of invasive lobular breast cancer. Cell, 163(2), 506–519. https://doi.org/10.1016/j.cell.2015.09.033
9. Clough, E., & Barrett, T. (2016). The Gene Expression Omnibus database. Methods in Molecular Biology (Clifton, N.J.), 1418, 93–110. https://doi.org/10.1007/978-1-4939-3578-9_5

10. Di Bisceglie, A. M. (2009). Hepatitis B and hepatocellular carcinoma. Hepatology (Baltimore, Md.), 49(5 Suppl), S56–S60. https://doi.org/10.1002/hep.22962
11. Duan, R., Jiang, L., Wang, T., Li, Z., Yu, X., Gao, Y., Jia, R., Fan, X., & Su, W. (2024). Aging-induced immune microenvironment remodeling fosters melanoma in male mice via γδ17-Neutrophil-CD8 axis. Nature Communications, 15(1), 10860. https://doi.org/10.1038/s41467-024-55164-3
12. Fane, M., & Weeraratna, A. T. (2020). How the ageing microenvironment influences tumour progression. Nature Reviews. Cancer, 20(2), 89–106. https://doi.org/10.1038/s41568-019-0222-9
13. Hai, H., Tamori, A., & Kawada, N. (2014). Role of hepatitis B virus DNA integration in human hepatocarcinogenesis. World Journal of Gastroenterology: WJG, 20(20), 6236–6243. https://doi.org/10.3748/wjg.v20.i20.6236
14. Hamdy, H., Yang, Y., Cheng, C., & Liu, Q. (2023). Identification of potential hub genes related to aflatoxin B1, liver fibrosis and hepatocellular carcinoma via integrated bioinformatics analysis. Biology, 12(2), 205. https://doi.org/10.3390/biology12020205
15. Hasan, A. N., Ahmad, M. W., Madar, I. H., Grace, B. L., & Hasan, T. N. (2015). An in silico analytical study of lung cancer and smokers datasets from gene expression omnibus (GEO) for prediction of differentially expressed genes. Bioinformation, 11(5), 229–235. https://doi.org/10.6026/97320630011229
16. Hasan, M. A. M., Maniruzzaman, M., & Shin, J. (2023). Differentially expressed discriminative genes and significant meta-hub genes based key genes identification for hepatocellular carcinoma using statistical machine learning. Scientific Reports, 13(1), 3771. https://doi.org/10.1038/s41598-023-30851-1
17. Hossen, M. A., Reza, M. S., Rana, M. M., Hossen, M. B., Shoaib, M., Mollah, M. N. H., & Han, C. (2024). Identification of most representative hub-genes for diagnosis, prognosis, and therapies of hepatocellular carcinoma. Chinese Clinical Oncology, 13(3), 32. https://doi.org/10.21037/cco-23-151
18. Jiang, N., Zhang, X., Qin, D., Yang, J., Wu, A., Wang, L., Sun, Y., Li, H., Shen, X., Lin, J., Kantawong, F., & Wu, J. (2021). Identification of core genes related to progression and prognosis of hepatocellular carcinoma and small-molecule drug predication. Frontiers in Genetics, 12, 608017. https://doi.org/10.3389/fgene.2021.608017
19. Kalaki, N. S., Ahmadzadeh, M., Mansouri, A., Saberiyan, M., & Karbalaie Niya, M. H. (2024). Identification of hub genes and pathways in hepatitis B virus-associated hepatocellular carcinoma: A comprehensive in silico study. Health Science Reports, 7(6), e2185. https://doi.org/10.1002/hsr2.2185
20. Llovet, J. M., Kelley, R. K., Villanueva, A., Singal, A. G., Pikarsky, E., Roayaie, S., Lencioni, R., Koike, K., Zucman-Rossi, J., & Finn, R. S. (2021). Hepatocellular carcinoma. Nature Reviews. Disease Primers, 7(1), 6. https://doi.org/10.1038/s41572-020-00240-3
21. Mi, N., Cao, J., Zhang, J., Fu, W., Huang, C., Gao, L., Yue, P., Bai, B., Lin, Y., Meng, W., & Li, X. (2020). Identification of hub genes involved in the occurrence and development of hepatocellular carcinoma via bioinformatics analysis. Oncology Letters, 20(2), 1695–1708. https://doi.org/10.3892/ol.2020.11752
22. Pan, Y., Chen, H., & Yu, J. (2020). Biomarkers in hepatocellular carcinoma: Current status and future perspectives. Biomedicines, 8(12), 576. https://doi.org/10.3390/biomedicines8120576
23. Parikh, N. D., Mehta, A. S., Singal, A. G., Block, T., Marrero, J. A., & Lok, A. S. (2020). Biomarkers for the early detection of hepatocellular carcinoma. Cancer Epidemiology, Biomarkers & Prevention: A Publication of the American Association for Cancer Research, Cosponsored by the American Society of Preventive Oncology, 29(12), 2495–2503. https://doi.org/10.1158/1055-9965.EPI-20-0005
24. Rao, V. S., Srinivas, K., Sujini, G. N., & Kumar, G. N. S. (2014). Protein-protein interaction detection: methods and analysis. International Journal of Proteomics, 2014, 147648. https://doi.org/10.1155/2014/147648
25. Rizzo, G. E. M., Cabibbo, G., & Craxì, A. (2022). Hepatitis B virus-associated hepatocellular carcinoma. Viruses, 14(5), 986. https://doi.org/10.3390/v14050986
26. Rosati, D., Palmieri, M., Brunelli, G., Morrione, A., Iannelli, F., Frullanti, E., & Giordano, A. (2024). Differential gene expression analysis pipelines and bioinformatic tools for the identification of specific biomarkers: A review. Computational and Structural Biotechnology Journal, 23, 1154–1168. https://doi.org/10.1016/j.csbj.2024.02.018
27. Shen, C., Jiang, X., Li, M., & Luo, Y. (2023). Hepatitis virus and hepatocellular carcinoma: Recent advances. Cancers, 15(2), 533. https://doi.org/10.3390/cancers15020533
28. Shen, Y., Huang, J., Jia, L., Zhang, C., & Xu, J. (2024). Bioinformatics and machine learning driven key genes screening for hepatocellular carcinoma. Biochemistry and Biophysics Reports, 37(101587), 101587.

https://doi.org/10.1016/j.bbrep.2023.101587

29. Thirumani, L., Helan, M., S, V., Jamal Mohamed, U., Vimal, S., & Madar, I. H. (2024). The molecular landscape of lung metastasis in primary head and neck squamous cell carcinomas. Cureus, 16(4), e57497. https://doi.org/10.7759/cureus.57497
30. Tomczak, K., Czerwińska, P., & Wiznerowicz, M. (2015). The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemporary Oncology (Poznan, Poland), 19(1A), A68–A77. https://doi.org/10.5114/wo.2014.47136
31. Xie, W., Wang, B., Wang, X., Hou, D., Su, H., & Huang, H. (2020). Nine hub genes related to the prognosis of HBV-positive hepatocellular carcinoma identified by protein interaction analysis. Annals of Translational Medicine, 8(7), 478. https://doi.org/10.21037/atm.2020.03.94
32. Xue, J.-M., Liu, Y., Wan, L.-H., & Zhu, Y.-X. (2020). Comprehensive analysis of differential gene expression to identify common gene signatures in multiple cancers. Medical Science Monitor: International Medical Journal of Experimental and Clinical Research, 26, e919953. https://doi.org/10.12659/MSM.919953
33. Yip, T. C.-F., Chan, H. L.-Y., Wong, V. W.-S., Tse, Y.-K., Lam, K. L.-Y., & Wong, G. L.-H. (2017). Impact of age and gender on risk of hepatocellular carcinoma after hepatitis B surface antigen seroclearance. Journal of Hepatology, 67(5), 902–908. https://doi.org/10.1016/j.jhep.2017.06.019
34. Zhang, J., Liu, X., Zhou, W., Lu, S., Wu, C., Wu, Z., Liu, R., Li, X., Wu, J., Liu, Y., Guo, S., Jia, S., Zhang, X., & Wang, M. (2021). Identification of key genes associated with the process of hepatitis B inflammation and cancer transformation by integrated bioinformatics analysis. Frontiers in Genetics, 12, 654517. https://doi.org/10.3389/fgene.2021.654517
35. Zhang, M., Chen, H., Liu, H., & Tang, H. (2024). The impact of integrated hepatitis B virus DNA on oncogenesis and antiviral therapy. Biomarker Research, 12(1), 84. https://doi.org/10.1186/s40364-024-00611-y
36. Zhao, L.-H., Liu, X., Yan, H.-X., Li, W.-Y., Zeng, X., Yang, Y., Zhao, J., Liu, S.-P., Zhuang, X.-H., Lin, C., Qin, C.-J., Zhao, Y., Pan, Z.-Y., Huang, G., Liu, H., Zhang, J., Wang, R.-Y., Yang, Y., Wen, W., … Wang, H.-Y. (2016). Genomic and oncogenic preference of HBV integration in hepatocellular carcinoma. Nature Communications, 7, 12992. https://doi.org/10.1038/ncomms12992
37. Zoulim, F., Chen, P.-J., Dandri, M., Kennedy, P. T., & Seeger, C. (2024). Hepatitis B virus DNA integration: Implications for diagnostics, therapy, and outcome. Journal of Hepatology, 81(6), 1087–1099. https://doi.org/10.1016/j.jhep.2024.06.037