Specialty Journal of Pharmacognosy, Phytochemistry, and Biotechnology
ISSN: 3062-441X
2024, Volume 4, Page No: 157-170

Available online at: www.galaxypub.co/page/journals

Galaxy Publication

Galaxy Publication

5-Fluorouracil Sensitivity-Based Molecular Subtyping of Hepatocellular Carcinoma Reveals Three Clinically Distinct Subgroups with Differential Prognosis

Diego Huaman[1], Mario Quispe[1], Cesar Flores[1]*, Luis Paredes[1]

[1]Department of Medicinal Plant Research, Faculty of Pharmacy, University of Cusco, Cusco, Peru.

***E-mail** ✉ cesar.flores.mpr@outlook.com

Received: 28 February 2024; **Revised:** 29 May 2024; **Accepted:** 01 June 2024

ABSTRACT

5-Fluorouracil (5-FU) is a commonly employed chemotherapeutic agent in multiple cancers, including hepatocellular carcinoma (HCC). Understanding why some HCC cases respond poorly or are resistant to 5-FU is essential for advancing precision oncology and optimizing treatment strategies. We applied Weighted Gene Co-expression Network Analysis (WGCNA) on gene expression data from the GDSC2 cancer cell line collection to detect 5-FU-related co-expression modules and hub genes. Based on these hub genes, HCC samples were classified into subgroups, and predictive models were developed using ConsensusClusterPlus combined with five machine learning algorithms. Additionally, the expression of key genes in the model was validated via quantitative reverse transcription-polymerase chain reaction (qRT-PCR).

WGCNA identified 19 distinct gene modules in the cancer cell lines, with the midnight blue module showing the strongest inverse association with 5-FU response. Within this module, 45 hub genes were selected. HCC patients were categorized into three subtypes: C1, C2, and C3. C1 had the poorest overall survival (OS) and was marked by a higher clinical grade and advanced T stage and stage, while C3 exhibited the most favorable OS. C2 had intermediate OS and displayed the lowest immune cell infiltration. From the 45 hub genes, five—TOMM40L, SNRPA, ILF3, CPSF6, and NUP205—were chosen to construct a prognostic regression model for HCC. qRT-PCR confirmed that these genes were markedly overexpressed in HCC tissue samples. Stratification of HCC according to 5-FU sensitivity aligns with prognostic differences and reflects heterogeneity in genomic features, immune infiltration, and signaling pathways. The derived 5-FU-related risk model may serve as a valuable tool for individualized prognosis monitoring in HCC patients.

Keywords: 5-fluorouracil, Hepatocellular carcinoma, Machine learning, Genomic variation, Immune infiltration, Prognosis

How to Cite This Article: Huaman D, Quispe M, Flores C, Paredes L. 5-Fluorouracil Sensitivity-Based Molecular Subtyping of Hepatocellular Carcinoma Reveals Three Clinically Distinct Subgroups with Differential Prognosis. Spec J Pharmacogn Phytochem Biotechnol. 2024;4:157-70. https://doi.org/10.51847/Aq24YvU5A8

Introduction

Hepatocellular carcinoma (HCC) represents a major global health challenge, with mortality rates approaching 91.6% relative to incidence [1]. High-risk individuals include those with hepatitis B or C infections, cirrhosis of any etiology, chronic alcohol use, non-alcoholic steatohepatitis, or family history of HCC, particularly men over 40 years old [2]. Surgical interventions such as hepatectomy or liver transplantation remain primary curative approaches [3] but are limited by patient suitability and donor organ availability [4]. Unfortunately, over 60% of HCC cases are detected at advanced stages, emphasizing the need for systemic therapies like sorafenib or lenvatinib [5]. Resistance to these drugs remains a significant obstacle, underscoring the importance of investigating molecular mechanisms of treatment failure.

5-FU, a synthetic fluorinated pyrimidine analog, acts as an anti-metabolite by mimicking essential cellular molecules. It disrupts nucleic acid synthesis by inhibiting thymidylate synthase and incorporating its metabolites into DNA and RNA, inducing cytotoxic effects [6, 7]. Since its FDA approval in 1962, 5-FU has been applied alone or in combination to treat diverse malignancies, including head and neck squamous cell carcinoma [8],

colorectal cancer [9], gastric cancer [10], metastatic breast cancer [11, 12], and HCC. Enhancing HCC sensitivity to 5-FU is critical [13], and identifying molecular determinants of resistance is essential for personalized therapy. Discovery of 5-FU predictive biomarkers and development of targeted therapies could improve patient outcomes [14].

In this study, we analyzed large-scale datasets to identify 5-FU-associated genes, classify HCC molecularly, and characterize tumor microenvironment heterogeneity. Candidate genes were further screened to develop a predictive risk model, potentially offering novel insights into mechanisms of 5-FU resistance in HCC.

Materials and Methods

Data sources

Clinical information and RNA-seq data for HCC were obtained from The Cancer Genome Atlas (TCGA) LIHC project (https://cancergenome.nih.gov), including 365 tumor samples and 50 matched adjacent non-tumor tissues. HCC microarray data (GSE14520) were retrieved from the Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/), and additional expression profiles were obtained from the HCCDB database (http://lifeome.net/database/hccdb/download.html). Pan-cancer cell line drug sensitivity data, along with genomic features, were downloaded from the Genomics of Drug Sensitivity in Cancer (GDSC, https://www.cancerrxgene.org/) resource [15].

Weighted gene co-expression network analysis (WGCNA)

WGCNA was applied to the expression profiles of cancer cell lines from GDSC2. Initial clustering of samples was performed to detect outliers and construct a gene co-expression network, from which modules were derived and correlated with external traits using the "WGCNA" R package [16]. The sampleTree function identified outlier cell lines. Gene expression matrices were used to calculate Pearson correlations between genes. Optimal soft threshold powers (β) were determined using pickSoftThreshold by evaluating scale independence and mean connectivity, while scaleFreePlot assessed whether the network approximated scale-free topology. Hierarchical clustering was performed with the hclust function. Gene modules were detected using cutreeDynamic (height = 0.25, deepSplit = 3), with a minimum module size of 30 genes. Modules showing relevance to 5-FU were prioritized, and module–trait relationships were assessed using IC50 values of 5-FU. Modules were merged via mergeCloseModules.

Unsupervised clustering of HCC

Differentially expressed genes (DEGs) were identified using the "limma" package [17], with FDR < 0.05 and |log2 fold change| > 1 as cutoffs. DEGs overlapping with 5-FU-related WGCNA modules were selected for consensus clustering. Using the ConsensusClusterPlus package [18], TCGA-LIHC samples were subsampled at 80% and partitioned into k clusters. Consensus matrices and cumulative distribution function (CDF) plots were generated to determine optimal clustering.

Single-nucleotide variant (SNV) and copy-number variant (CNV) analyses

Genomic alterations, including SNVs, indels, and CNVs, were analyzed. SNVs were evaluated from TCGA MAF files using the "maftools" R package [19], which also produced oncoplots. CNV analysis was performed with GISTIC 2.0, which identifies genomic regions with frequent alterations, weighting high-amplitude events such as high-level gains or homozygous deletions [20]. Peak regions with maximal aberration frequency and amplitude were defined for each significant locus.

Immune cell infiltration analysis

Immune and stromal cell content was inferred using the ESTIMATE algorithm based on TCGA-LIHC transcriptional profiles [21]. Multiple approaches were applied to quantify immune infiltration: CIBERSORT evaluated relative proportions of 22 immune cell types within samples [22], MCPcounter estimated absolute abundance of cell populations across samples [23], and TIMER considered tissue specificity to quantify six immune cell types, allowing association analyses between infiltration levels and HCC clusters [24].

Development of a risk stratification model using machine learning

Univariate Cox regression was performed on DEGs intersecting with 5-FU-related modules to identify prognostic genes. Five machine learning algorithms—Gradient Boosting Machine (GBM), LASSO regression, Support Vector Machine (SVM), Decision Trees, and Random Forest—were employed for feature selection. Genes identified by all five methods were subjected to stepwise multiple linear regression to construct a final prognostic model, allowing calculation of risk scores for patients in different HCC cohorts.

Nomogram construction

We combined clinical variables—including age, sex, T stage, overall stage, tumor grade, and RiskScore—to perform both univariate and multivariate Cox regression analyses to pinpoint independent predictors of HCC prognosis. These identified factors were then utilized to construct a nomogram for estimating patient survival probabilities. Calibration curves were plotted to compare predicted versus actual survival, confirming the accuracy of the nomogram. In addition, decision curve analysis (DCA) was conducted to evaluate the potential clinical benefit and predictive value of both the nomogram and RiskScore.

Cell culture and transient transfection

The hepatocellular carcinoma cell lines Hep3B2.1-7 and Huh-7 were obtained from COBIOER (Nanjing, China) and maintained in DMEM/F12 medium supplemented with 10% fetal bovine serum (Gibco, Thermo Fisher, USA). Normal human liver epithelial THLE-3 cells were procured from ATCC (Manassas, VA, USA) and cultured in BEGM medium (Lonza, Walkersville). All cells were incubated at 37°C with 5% CO2 in a humidified environment.

Quantitative reverse transcription polymerase chain reaction (qRT-PCR)

Total RNA was extracted from Hep3B2.1-7, Huh-7, and THLE-3 cells using TRIzol reagent (Thermo Fisher, USA). For qRT-PCR, 2 μg of RNA was converted into cDNA and amplified using FastStart Universal SYBR Green Master Mix (Roche, Switzerland) on a LightCycler 480 system (Roche, USA). Each reaction (20 μl total volume) contained 2 μl of cDNA, 10 μl SYBR Green mix, 0.5 μl of each primer, and nuclease-free water. PCR cycling included initial denaturation at 95°C for 30 s, followed by 45 cycles of 94°C for 15 s, 56°C for 30 s, and 72°C for 20 s. Each sample was analyzed in triplicate. Relative expression was calculated using the 2^−ΔΔCT method, normalized to GAPDH, and compared to expression in normal liver cells.

Statistical analysis

All analyses were performed using R software. Differences in clinical features among subgroups were evaluated using the chi-square test. Kaplan–Meier curves were generated to illustrate survival distributions, and statistical significance was assessed using the log-rank test. Time-dependent ROC curves and AUC values were calculated via the timeROC package to assess the predictive performance of the risk stratification model. A p-value < 0.05 was considered statistically significant. Graphical representation used the following symbols: ns = not significant, * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$.

Results and Discussion

Identification of the 5-FU-associated gene module

We evaluated 5-FU sensitivity across HCC cell lines. Hep3B2.1-7 cells displayed the lowest IC50, indicating maximal sensitivity, whereas Huh-7 cells exhibited the highest IC50, reflecting strong resistance **(Figure 1a)**. All GDSC2 cell line samples were clustered **(Figure 1b)**. The soft-threshold power selected to satisfy scale-free network criteria was 6, achieving R^2 = 0.86, and the mean connectivity approached zero **(Figure 1c)**. Nineteen co-expression modules were identified through hierarchical clustering of all genes **(Figure 1d)**. Among these, correlation analysis with 5-FU IC50 values indicated that the midnight blue module had the strongest significant negative association **(Figure 1e)**. Functional enrichment of the midnight blue module genes revealed involvement in mRNA processing regulation, polyadenylation control, and positive telomere capping. These genes encoded proteins that were part of the transcription elongation factor complex and DNA polymerase III complex **(Figure 1f)**.

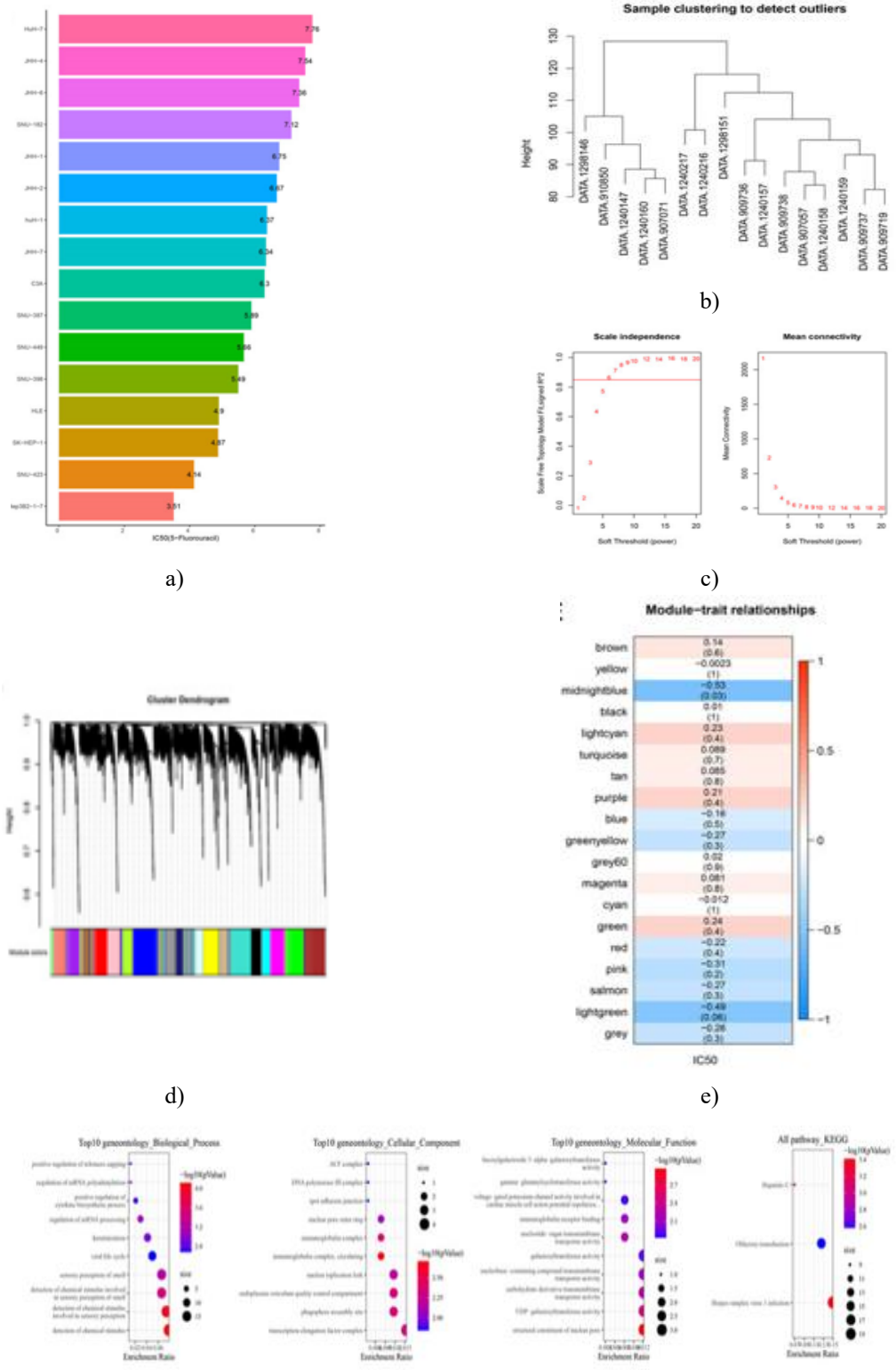


Figure 1. Identification of 5-fluorouracil-associated gene module

(a) Bar chart illustrating 5-fluorouracil sensitivity across various hepatocellular carcinoma cell lines. (b) Dendrogram of cell line samples from GDSC2. (c) Plot showing average connectivity relative to the scale-free fit index across different soft-threshold values. (d) Hierarchical clustering of all genes in hepatoma cell lines from GDSC2. (e) Correlation analysis between 19 gene modules and 5-fluorouracil IC50; the upper numbers indicate correlation coefficients, while the lower numbers represent p-values. (f) GO terms and KEGG pathways enriched in the midnight blue module.

Classification of HCC based on hub genes from the midnight blue module

Expression differences between normal liver tissues and HCC samples (log2 [TPM+1]) from TCGA-LIHC were analyzed. A total of 2,356 genes were significantly upregulated (log2 [Fold Change] > 1, FDR < 0.05) and 462 genes were downregulated (log2 [Fold Change] < −1, FDR < 0.05) **(Figure 2a)**. Among the midnight blue module genes, 35 were upregulated, and 8 were downregulated **(Figure 2b)**. Using these 43 hub genes, TCGA-LIHC samples were clustered. The cumulative distribution function (CDF) suggested an optimal k-value of 3 **(Figure 2c)**. The consensus matrix confirmed the clustering into three groups **(Figure 2d)**. Kaplan–Meier survival curves demonstrated significant differences in overall survival (OS) among the clusters in TCGA-LIHC, HCCB18, and GSE145203 cohorts. Cluster C3 had the best prognosis, while C1 showed the shortest OS, and C2 had intermediate survival **(Figures 2e–2g)**. A heatmap of the 43 genes revealed higher expression in C1 compared with C2 and C3 **(Figure 2h)**.

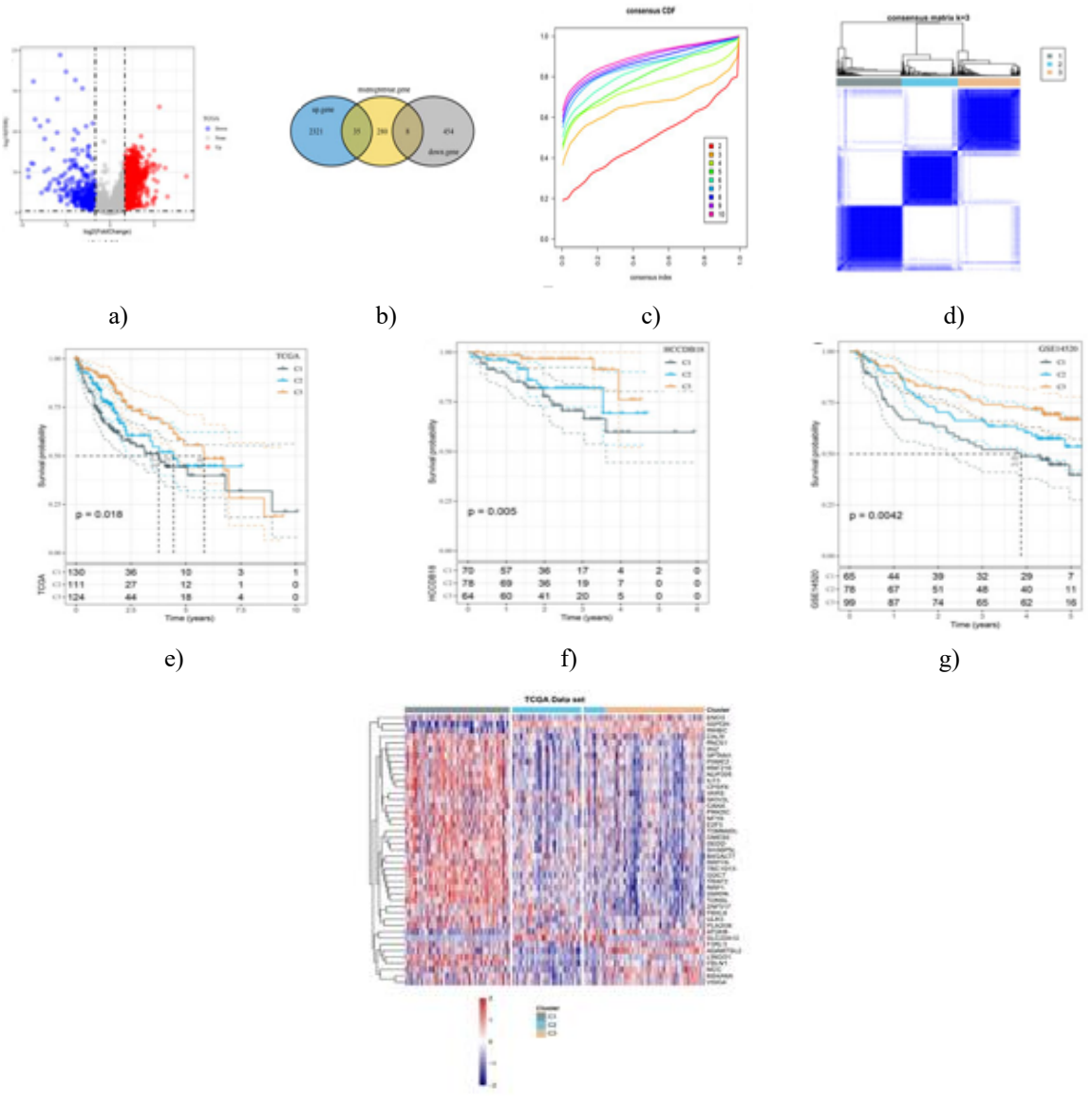
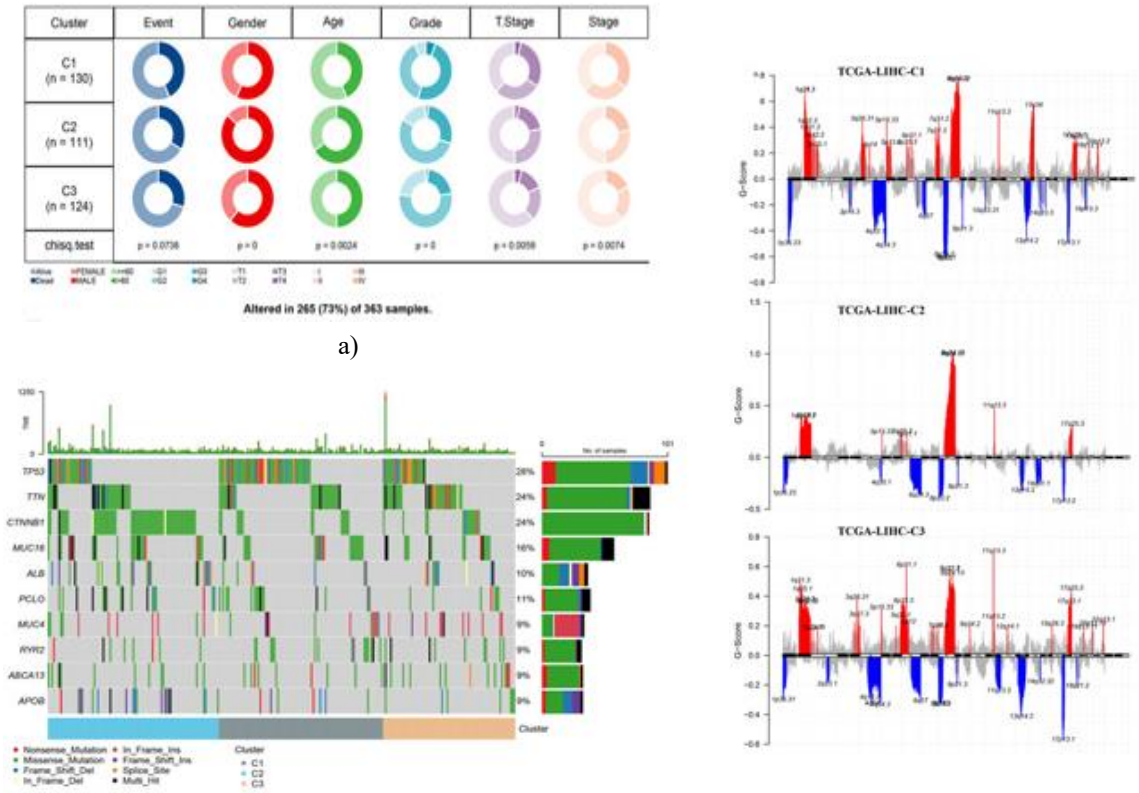


Figure 2. Classification of HCC by midnight blue hub genes

(a) Differential expression analysis (log2 [TPM+1]) between normal and HCC tissues in TCGA-LIHC. (b) Overlap analysis between DEGs and the midnight blue module to identify hub genes. (c) CDF plot showing consensus distributions for different k-values. (d) Consensus matrix showing k = 3 clustering; blue indicates sample similarity. (e–g) Kaplan–Meier survival curves for clusters in TCGA-LIHC, HCCB18, and GSE145203 cohorts. (h) Heatmap of 43 hub genes expression across clusters.

Clinical and genomic characteristics of molecular subtypes

The three molecular subtypes (C1, C2, and C3) were analyzed for genomic alterations. TP53 mutations were significantly enriched in C1 relative to C2 and C3. C2 had the highest frequency of CTNNB1 mutations, while TTN mutations were most common in C3 **(Figure 3a)**. Chromosomal CNV analysis using Manhattan plots showed fewer high-level amplifications and deletions in C2 compared with C1 and C3 **(Figure 3b)**. Clinical feature comparison revealed a predominance of male patients across subtypes. Age, sex, tumor grade, T stage, and overall stage differed significantly among clusters. C2 contained the highest proportion of males and patients older than 60 years, whereas C1, which had the shortest OS, was associated with higher tumor grade and advanced T stage and overall stage **(Figure 3c)**.



a) b) c)

Figure 3. Clinical and genomic features of molecular subtypes

(a) Waterfall plot of somatic mutations across three subtypes. (b) Manhattan plot showing CNV distribution at the chromosomal level. (c) Comparison of clinical characteristics among clusters. Red indicates the q-arm and blue indicates the p-arm of chromosomes.

Immune infiltration patterns across three clusters

Using ESTIMATE, stromal, immune, and total ESTIMATE scores were calculated, revealing significant differences among clusters, with the lowest levels observed in C2 **(Figure 4a)**. Of the 22 immune cell types measured by CIBERSORT, 15 showed significant differences across clusters **(Figure 4b)**. Analysis of 28 tumor-infiltrating lymphocyte (TIL) subpopulations indicated that memory, immunosuppressive (Tregs and MDSCs), and cytotoxic (CD8 T, NK, and NKT) cells were differentially infiltrated, with C2 showing the lowest levels **(Figure 4c)**. MCPcounter and TIMER results also showed reduced infiltration of CD4 T cells, total T cells, B cells, macrophages, neutrophils, CD8 T cells, endothelial cells, dendritic cells, and fibroblasts in C2 compared with C1 and C3 **(Figures 4d and 4e)**. A heatmap of pathway enrichment revealed that most metabolic pathways, including linoleic acid, tyrosine, phenylalanine, and pyruvate metabolism, were less enriched in C1 **(Figure 4f)**.

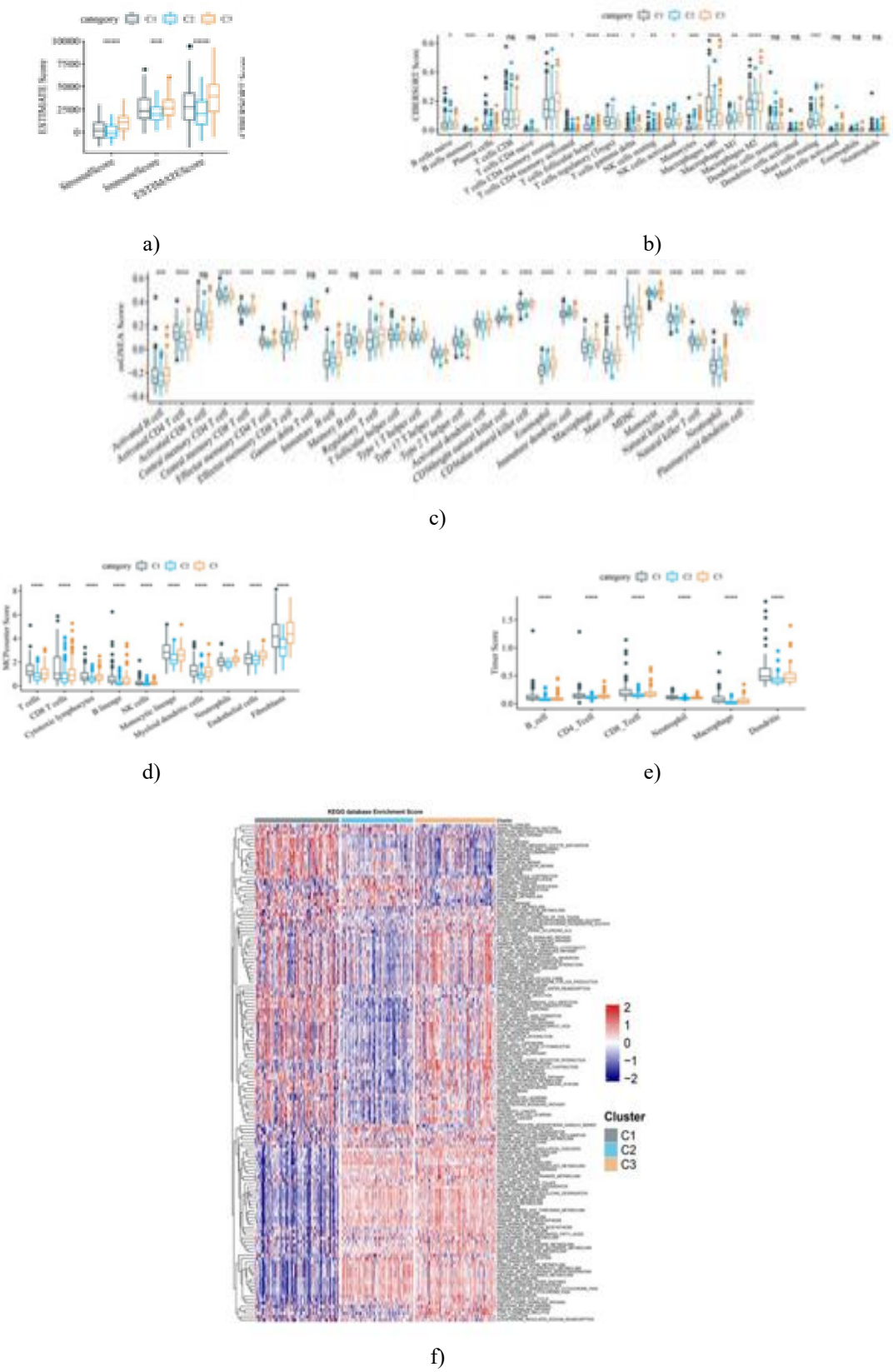


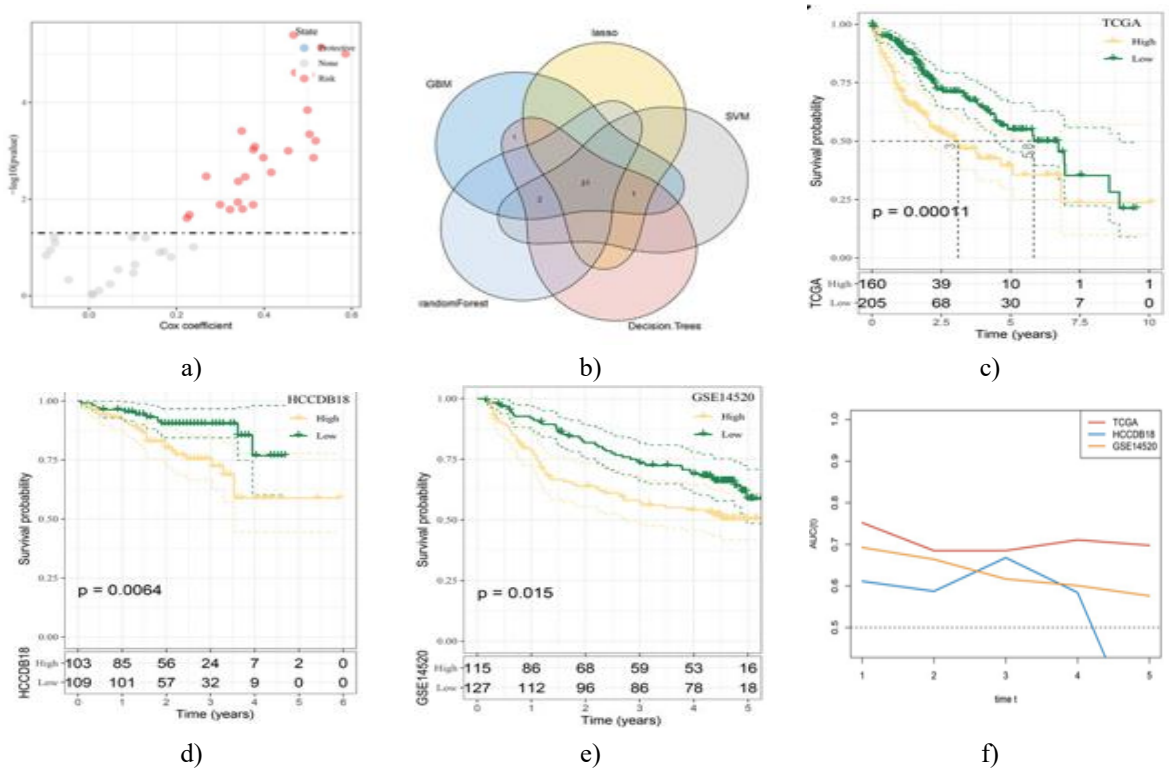
Figure 4. Immune infiltration patterns across clusters
(a) Stromal, immune, and ESTIMATE scores for each cluster. (b–e) Immune cell infiltration differences were assessed using CIBERSORT, ssGSEA, MCPcounter, and TIMER. (f) Heatmap showing pathway enrichment differences among clusters.

Identification and validation of candidate genes for the risk model in the midnight blue module

Within the midnight blue module, 43 hub genes were initially recognized. Univariate Cox regression analysis further filtered 25 genes associated with HCC prognosis **(Figure 5a)**. To refine candidate genes, five machine learning approaches—LASSO, GBM, SVM, Random Forest, and Decision Tree—were applied, resulting in 21 overlapping genes across all models **(Figure 5b)**. Stepwise regression then selected five genes suitable for constructing the final risk model: TOMM40L, SNRPA, ILF3, CPSF6, and NUP205. Multivariate Cox regression generated the risk coefficients and the resulting RiskScore formula:

$$\text{RiskScore} = 0.293 \times \text{TOMM40L} + 0.558 \times \text{SNRPA} - 0.823 \times \text{ILF3} + 0.493 \times \text{CPSF6} + 0.464 \times \text{NUP205} \quad (1)$$

RiskScore was calculated for the TCGA-LIHC cohort and two independent validation datasets (HCCDB18 and GSE14520). Across all datasets, higher RiskScores were associated with shorter overall survival (OS) **(Figures 5c–5e)**. The model also displayed stable predictive performance for 1–5-year OS in TCGA-LIHC and GSE14520 **(Figure 5f)**.



a) b) c)

d) e) f)

Figure 5. Gene selection and risk model validation in the midnight blue module

(a) Volcano plot from univariate Cox regression for 43 hub genes. (b) Overlap of genes selected by LASSO, GBM, SVM, Random Forest, and Decision Tree. (c–e) Kaplan–Meier survival curves for TCGA-LIHC, HCCDB18, and GSE14520 stratified by RiskScore. (f) ROC curves assessing 1–5-year OS prediction.

Prognostic value of the risk model and clinical feature associations

A heatmap of clinical variables versus RiskScore revealed significant differences between high- and low-risk groups, including molecular subtype, survival status, T stage, overall stage, and tumor grade. The high-risk group exhibited a higher C1 proportion, elevated mortality, advanced T stage and stage, and higher tumor grades, whereas the low-risk group showed milder clinical features **(Figure 6a)**. Univariate Cox regression confirmed that RiskScore, T stage, and stage were associated with prognosis. Multivariate analysis established RiskScore as an independent prognostic indicator **(Figures 6b and 6c)**. Using RiskScore alongside T stage and stage, a nomogram was constructed to predict 1-, 3-, and 5-year clinical outcomes. Calibration curves demonstrated close alignment between predicted and actual outcomes, while decision curve analysis confirmed the clinical applicability of the nomogram and RiskScore. RiskScore was significantly higher in advanced T3–T4 stage, stage III–IV, and G3–G4 samples compared to early-stage T1–T2, stage I–II, and G1–G2 samples **(Figures 6d–6f)**.

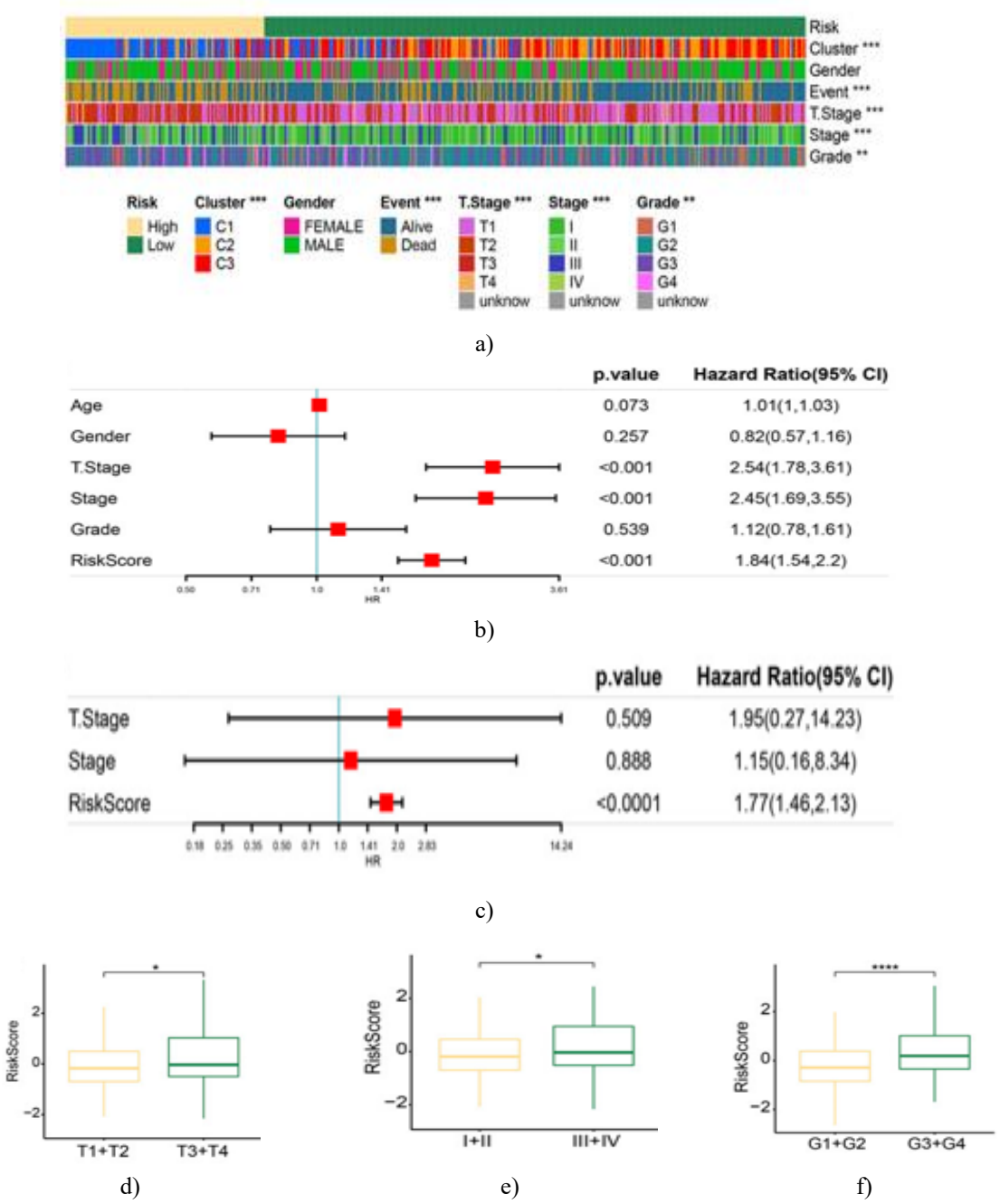
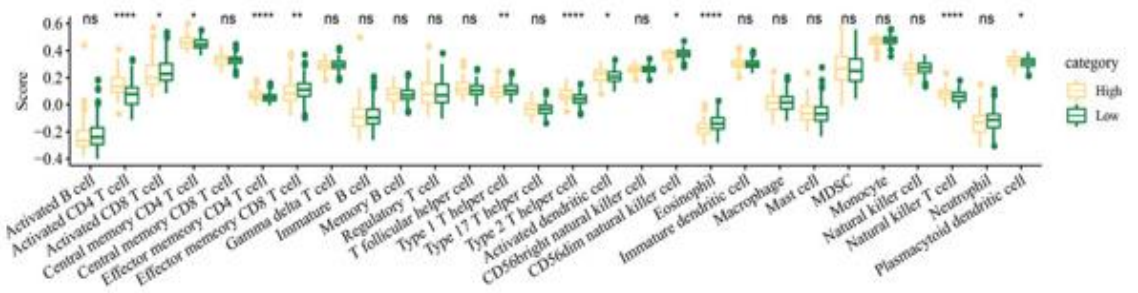


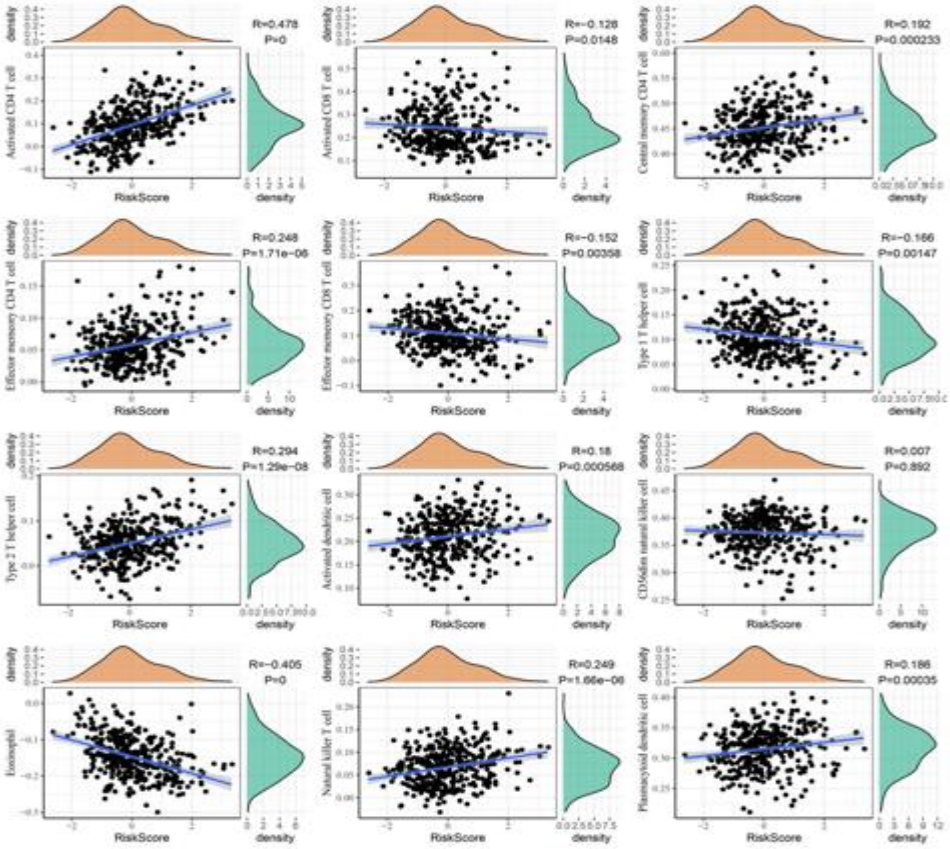
Figure 6. Clinical and prognostic relevance of the risk model in HCC
(a) Heatmap linking RiskScore to clinical features. (b) Univariate Cox regression for RiskScore and clinical variables. (c) Multivariate Cox regression confirms RiskScore as an independent predictor. (d–f) RiskScore differences stratified by T stage, stage, and tumor grade.

Association between RiskScore and immune infiltration

Immune cell infiltration analysis demonstrated significant differences between high- and low-risk groups for 12 immune cell types, including type 2 T helper cells, central memory CD4 T cells, type 1 T helper cells, plasmacytoid dendritic cells, effector memory CD4 T cells, activated CD4 T cells, activated CD8 T cells, eosinophils, NK T cells, CD56^dim NK cells, activated dendritic cells, and effector memory CD8 T cells **(Figure 7a)**. Correlation analysis using ssGSEA indicated strong associations for 11 cell types with RiskScore. Specifically, effector memory CD4 T cells, activated CD4 T cells, NK T cells, plasmacytoid dendritic cells, type 2 T helper cells, central memory CD4 T cells, and activated dendritic cells were positively correlated, whereas activated CD8 T cells, effector memory CD8 T cells, type 1 T helper cells, and eosinophils were negatively correlated. CD56^dim NK cells showed negligible correlation with RiskScore **(Figure 7b)**.



a)



b)

Figure 7. Association between RiskScore and immune cell infiltration

(a) Immune infiltration levels were assessed based on stratification by RiskScore. (b) Correlation between ssGSEA scores of RiskScore and immune cell populations.

PCR validation of RiskScore

To confirm the robustness of the RiskScore, we evaluated the expression levels of the five model genes using qRT-PCR. Results confirmed the reliability of the RiskScore, showing significant upregulation of TOMM40L, SNRPA, ILF3, CPSF6, and NUP205 in HCC cell lines Hep3B2.1-7 and Huh-7 relative to normal liver epithelial cells THLE-3 **(Figures 8a–8e)**.

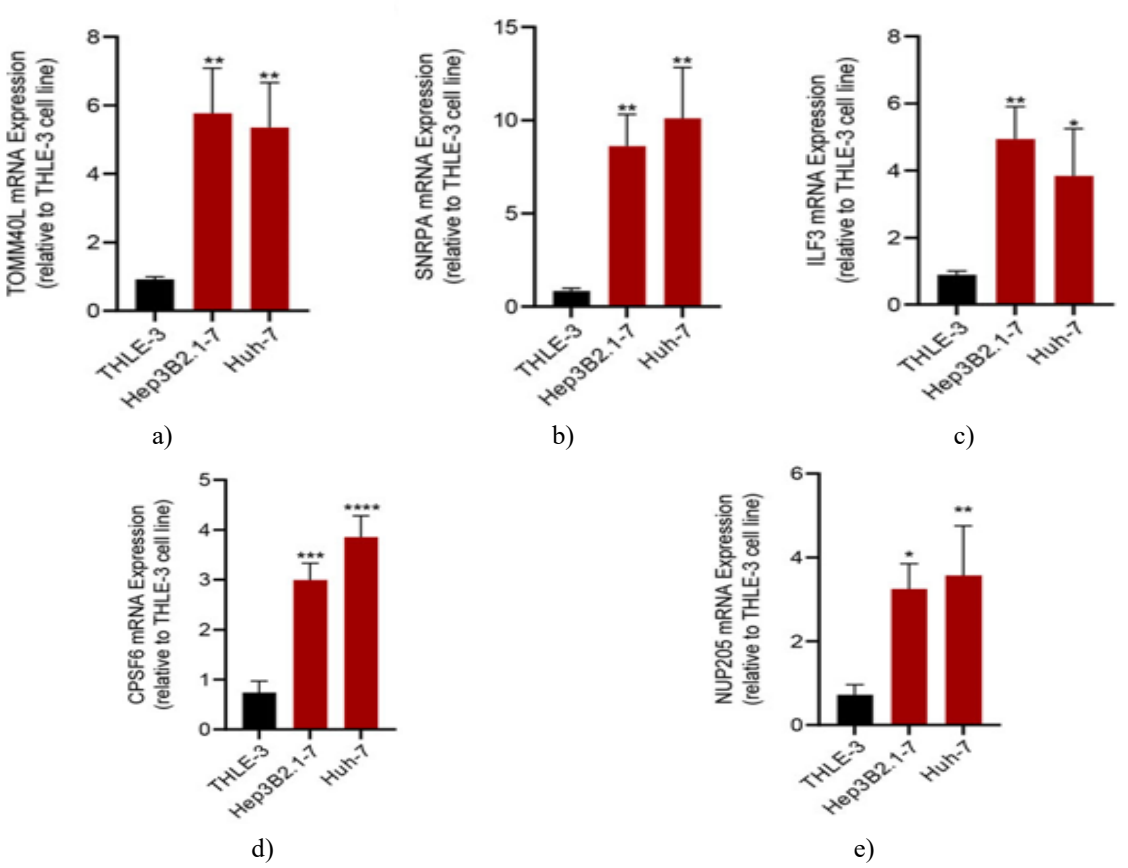


Figure 8. qRT-PCR validation of the five RiskScore genes
(a) TOMM40L; (b) SNRPA; (c) ILF3; (d) CPSF6; (e) NUP205

Large-scale cancer cell line pharmacogenomics databases provide comprehensive molecular characterization and detailed drug response profiles, offering a valuable resource to explore mechanisms underlying chemosensitivity [25]. In this study, we examined the heterogeneity of HCC in the context of 5-FU sensitivity using expression profiles from a large public dataset alongside drug response information. WGCNA identified the midnight blue module as most associated with 5-FU sensitivity, containing 43 hub genes. Based on these genes, three HCC subgroups were defined, reflecting differences in genomic alterations, tumor microenvironment (TME), and pathological pathway activity.

5-FU remains a widely used chemotherapeutic for tumors, exerting cytotoxicity by interfering with DNA and RNA synthesis, disrupting cell cycle progression, and inducing apoptosis [6, 14, 26]. Disruption of the tumor cell cycle is critical for halting proliferation and potentially restoring immune surveillance [27]. GO and KEGG analyses indicated that midnight blue module genes are linked to mRNA processing, regulation of polyadenylation, positive regulation of telomere capping, and components of the transcription elongation factor and DNA polymerase III complexes. By impairing homologous recombination repair, 5-FU promotes DNA damage and inhibits tumor cell growth [28]. These biological processes may underlie HCC sensitivity to 5-FU.

Regarding clinical outcomes, C3 displayed the best survival, while C1 had the poorest OS. C2 was notable for a higher proportion of male patients over 60 years old. The C1 subtype, with the shortest OS, was enriched in higher clinical grade, advanced T stage, and overall stage, and showed reduced activity in most metabolic pathways. Genomic profiling revealed TP53 mutations dominated C1, CTNNB1 mutations were highest in C2, and TTN mutations predominated in C3. This suggests that C1 represents a TP53-driven tumor subtype, C2 is CTNNB1-driven, and C3 is TTN-driven. TP53 and CTNNB1 mutations are common in HCC, with TP53 alterations linked to cell cycle dysregulation and impaired DNA repair [29]. Specific TP53 mutations (e.g., R249S) increase HCC risk, and TP53 loss enhances tumor cell viability and poor prognosis [30]. C3, although enriched with TTN mutations, may have a distinct mutational profile; previous work indicated that TTN mutations in blood samples predict poor outcomes in HCC [31]. Differences in tissue versus blood sampling may explain discrepancies.

Overall, our survival analysis aligns with molecular and clinical characteristics: C1 predicts poor outcomes, C2 intermediate, and C3 a favorable prognosis. HCC heterogeneity, both in genomic composition and mutational patterns, remains a key challenge [32]. TP53-enriched patients are likely to have worse survival, while CTNNB1-enriched C2 patients show active metabolic reprogramming, suggesting potential for targeting glycolytic pathways. These insights could inform precision treatment strategies tailored to C1 and C2 subtypes, improving therapeutic outcomes.

Personalized therapy for HCC patients has gained increasing attention and is being progressively applied in clinical practice. The establishment of risk prediction models constitutes a critical step toward individualized monitoring of HCC. Despite the publication of numerous risk models, only a few are routinely implemented in clinical settings to guide HCC surveillance and management decisions [33]. In this work, five out of the 43 hub genes from the midnight blue module were selected to construct a risk regression model. This model demonstrated independence and strong predictive capability for HCC prognosis while also reflecting relevant clinical features. Some of these genes have recognized pathological roles and regulatory mechanisms in tumorigenesis. For instance, SNRPA acts as a splicing factor linked to microvascular invasion and enhances HCC metastasis via the circSEC62/miR-625–5p-mediated activation of the NOTCH1/Snail pathway [34]. ILF3 is overexpressed in primary colorectal cancer and promotes tumor proliferation by stabilizing SGOC gene mRNA [35]. CPSF6 is upregulated in HCC and drives metabolic reprogramming in hepatocytes through NQO1 alternative polyadenylation [36]. While these genes have individually been implicated in cancer, integrating them into a single risk model represents a novel approach and provides a potential prognostic tool for HCC.

Conclusion

In conclusion, this study stratified HCC into subtypes based on 5-FU sensitivity. The classification aligned with observed differences in patient prognosis and captured the heterogeneity of genomic alterations, tumor microenvironment, and pathological signaling pathways. Additionally, we established an independent risk regression model incorporating five 5-FU-associated genes, offering a valuable resource for individualized monitoring and prognostic assessment of HCC patients.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. Villarruel-Melquiades F, Mendoza-Garrido ME, Garcia-Cuellar CM, Sanchez-Perez Y, Perez-Carreon JI, Camacho J. Current and novel approaches in the pharmacological treatment of hepatocellular carcinoma. World J Gastroenterol. 2023;29(17):2571-99. doi:10.3748/wjg.v29.i17.2571
2. Xie D, Shi J, Zhou J, Fan J, Gao Q. Clinical practice guidelines and real-life practice in hepatocellular carcinoma: a Chinese perspective. Clin Mol Hepatol. 2023;29(2):206-16. doi:10.3350/cmh.2022.0402
3. Kawaguchi Y, Honda G, Endo I, Cherqui D, Kokudo N. Current technical issues for surgery of primary liver cancer. Liver Cancer. 2017;6(1):51-8. doi:10.1159/000449345
4. Koza A, Bhogal RH, Fotiadis N, Mavroeidis VK. The role of ablative techniques in the management of hepatocellular carcinoma: indications and outcomes. Biomedicines. 2023;11(4):1062. doi:10.3390/biomedicines11041062
5. Yang C, Zhang H, Zhang L, Zhu AX, Bernards R, Qin W, et al. Evolving therapeutic landscape of advanced hepatocellular carcinoma. Nat Rev Gastroenterol Hepatol. 2023;20(4):203-22. doi:10.1038/s41575-022-00704-9
6. Blondy S, David V, Verdier M, Mathonnet M, Perraud A, Christou N. 5-Fluorouracil resistance mechanisms in colorectal cancer: from classical pathways to promising processes. Cancer Sci. 2020;111(9):3142-54. doi:10.1111/cas.14532

7. Mafi A, Rezaee M, Hedayati N, Hogan SD, Reiter RJ, Aarabi MH, et al. Melatonin and 5-fluorouracil combination chemotherapy: opportunities and efficacy in cancer therapy. Cell Commun Signal. 2023;21(1):33. doi:10.1186/s12964-023-01047-x
8. Yamauchi M, Minesaki A, Ishida T, Sato Y, Okamura S, Shuto H, et al. Induction chemotherapy with 5-fluorouracil, cisplatin, and cetuximab in advanced head and neck squamous cell carcinoma. In Vivo. 2023;37(3):1275-80. doi:10.21873/invivo.13205
9. Wosiak A, Szmajda-Krygier D, Pietrzak J, Boncela J, Balcerczak E. Assessment of the influence of 5-fluorouracil on SMAD4 and TGFB1 gene expression, apoptosis induction and DNA damage in human cell lines. Bioengineering (Basel). 2023;10(5):570. doi:10.3390/bioengineering10050570
10. Kang BW, Kim JG, Kwon OK, Chung HY, Yu W. Non-platinum-based chemotherapy for treatment of advanced gastric cancer: 5-fluorouracil, taxanes, and irinotecan. World J Gastroenterol. 2014;20(18):5396-402. doi:10.3748/wjg.v20.i18.5396
11. Karapetis CS, Patterson WK, Pittman KB, Kotasek D, Sage RE. Treatment of metastatic breast cancer with continuous infusional 5-fluorouracil. Aust N Z J Med. 1999;29(4):517-22. doi:10.1111/j.1445-5994.1999.tb00753.x
12. Holmes FA, Hellerstedt BA, Pippen JE, Vukelja SJ, Collea RP, Kocs DM, et al. Five-year results of a phase II trial of preoperative 5-fluorouracil, epirubicin, cyclophosphamide followed by docetaxel with capecitabine (wTX) (with trastuzumab in HER2-positive patients) for patients with stage II or III breast cancer. Cancer Med. 2018;7(6):2288-98. doi:10.1002/cam4.1472
13. Hu Z, Lv G, Li Y, Li E, Li H, Zhou Q, et al. Enhancement of anti-tumor effects of 5-fluorouracil on hepatocellular carcinoma by low-intensity ultrasound. J Exp Clin Cancer Res. 2016;35(1):71. doi:10.1186/s13046-016-0349-4
14. Vodenkova S, Buchler T, Cervena K, Veskrnova V, Vodicka P, Vymetalkova V. 5-Fluorouracil and other fluoropyrimidines in colorectal cancer: past, present and future. Pharmacol Ther. 2020;206(1):107447. doi:10.1016/j.pharmthera.2019.107447
15. Yang W, Soares J, Greninger P, Edelman EJ, Lightfoot H, Forbes S, et al. Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013;41(D1):D955-61. doi:10.1093/nar/gks1111
16. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9(1):559. doi:10.1186/1471-2105-9-559
17. 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(7):e47. doi:10.1093/nar/gkv007
18. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics. 2010;26(12):1572-3. doi:10.1093/bioinformatics/btq170
19. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. Genome Res. 2018;28(11):1747-56. doi:10.1101/gr.239244.118
20. Beroukhim R, Getz G, Nghiemphu L, Barretina J, Hsueh T, Linhart D, et al. Assessing the significance of chromosomal aberrations in cancer: methodology and application to glioma. Proc Natl Acad Sci U S A. 2007;104(50):20007-12. doi:10.1073/pnas.0710052104
21. Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4(1):2612. doi:10.1038/ncomms3612
22. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. Nat Methods. 2015;12(5):453-7. doi:10.1038/nmeth.3337
23. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;17(1):218. doi:10.1186/s13059-016-1070-5
24. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, et al. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. Cancer Res. 2017;77(21):e108-10. doi:10.1158/0008-5472.CAN-17-0307
25. Kusch N, Schuppert A. Two-step multi-omics modelling of drug sensitivity in cancer cell lines to identify driving mechanisms. PLoS One. 2020;15(11):e0238961. doi:10.1371/journal.pone.0238961

26. Sethy C, Kundu CN. 5-Fluorouracil resistance and the new strategy to enhance the sensitivity against cancer: implication of DNA repair inhibition. Biomed Pharmacother. 2021;137(1):111285. doi:10.1016/j.biopha.2021.111285
27. Liu J, Peng Y, Wei W. Cell cycle on the crossroad of tumorigenesis and cancer therapy. Trends Cell Biol. 2022;32(1):30-44. doi:10.1016/j.tcb.2021.07.001
28. Srinivas US, Dyczkowski J, Beissbarth T, Gaedcke J, Mansour WY, Borgmann K, et al. 5-Fluorouracil sensitizes colorectal tumor cells towards double-stranded DNA breaks by interfering with homologous recombination repair. Oncotarget. 2015;6(14):12574-86. doi:10.18632/oncotarget.3728
29. Gao Q, Zhu H, Dong L, Shi W, Chen R, Song Z, et al. Integrated proteogenomic characterization of HBV-related hepatocellular carcinoma. Cell. 2019;179(5):1240-54.e22. doi:10.1016/j.cell.2019.10.038
30. Lam YK, Yu J, Huang H, Ding X, Wong AM, Leung HH, et al. TP53 R249S mutation in hepatic organoids captures the predisposing cancer risk. Hepatology. 2023;78(3):727-40. doi:10.1002/hep.32802
31. Kunadirek P, Chuaypen N, Jenjaroenpun P, Wongsurawat T, Pinjaroen N, Sirichindakul P, et al. Cell-free DNA analysis by whole-exome sequencing for hepatocellular carcinoma: a pilot study in Thailand. Cancers (Basel). 2021;13(9):2229. doi:10.3390/cancers13092229
32. Jeng KS, Chang CF, Jeng WJ, Sheen IS, Jeng CJ. Heterogeneity of hepatocellular carcinoma contributes to cancer progression. Crit Rev Oncol Hematol. 2015;94(3):337-47. doi:10.1016/j.critrevonc.2015.01.009
33. Innes H, Nahon P. Statistical perspectives on using hepatocellular carcinoma risk models to inform surveillance decisions. J Hepatol. 2023;79(5):1113-21. doi:10.1016/j.jhep.2023.05.005
34. Mo Z, Li R, Cao C, Li Y, Zheng S, Wu R, et al. Splicing factor SNRPA associated with microvascular invasion promotes hepatocellular carcinoma metastasis through activating NOTCH1/Snail pathway and is mediated by circSEC62/miR-625-5p axis. Environ Toxicol. 2023;38(5):1022-37. doi:10.1002/tox.23745
35. Li K, Wu JL, Qin B, Fan Z, Tang Q, Lu W, et al. ILF3 is a substrate of SPOP for regulating serine biosynthesis in colorectal cancer. Cell Res. 2020;30(2):163-78. doi:10.1038/s41422-019-0257-1
36. Tan S, Zhang M, Shi X, Ding K, Zhao Q, Guo Q, et al. CPSF6 links alternative polyadenylation to metabolism adaption in hepatocellular carcinoma progression. J Exp Clin Cancer Res. 2021;40(1):85. doi:10.1186/s13046-021-01884-z