Journal of Hepatocellular Carcinoma

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Integrated Machine Learning and Multi-Omics Identifies a Novel Molecular Signature for Improving the Prognosis of Hepatocellular Carcinoma

Zhongxu Wu[1,*], Jianguang Xiong[2,*], Qisheng Liu[2,*], Chengdang Wang[1,3], Dan Li[4], Liuliu Wei[1], Jian Ding[1,3]

[1]Department of Gastroenterology, the First Affiliated Hospital of Fujian Medical University, Fuzhou, 350004, People's Republic of China; [2]Department of Gastroenterology, Xianning Central Hospital, Xianning, Hubei, 437000, People's Republic of China; [3]Department of Gastroenterology, National Regional Medical Center, Binhai Campus of the First Affiliated Hospital, Fujian Medical University, Fuzhou, 350200, People's Republic of China; [4]Department of Gastroenterology, Fujian Medical University Union Hospital, Fuzhou, 350001, People's Republic of China

*These authors contributed equally to this work

Correspondence: Jian Ding, Email docdingjian@163.com

Journal of Hepatocellular Carcinoma downloaded from https://www.dovepress.com/
For personal use only.

Background: Hepatocellular carcinoma (HCC) exhibits significant molecular heterogeneity and complex immune microenvironment, which to some extent limits the accuracy of prognosis assessment and the formulation of individualized treatment strategies. This study aims to identify immune-derived molecular signatures based on multi-omics data and machine learning methods for the prognosis prediction and risk stratification of HCC.

Methods: Based on weighted gene co-expression network analysis(WGCNA) and differential gene analysis,immune-derived molecular signature (IDMS) were screened in both single-cell and bulk transcriptomes. Prognostic model was constructed by multi-machine learning approachs. Subsequently, we investigated the differences in mutations, biological functions, and immune cell infiltration within the tumor microenvironment between the high- and low-risk groups.In addition, we comprehensively analyzed the drug sensitivity of IDMS and predicted potential drugs.

Results: We identified seven hub genes at the single-cell and bulk transcriptome levels. Based on multiple machine learning, we constructed a prognostic model that demonstrated excellent performance in predicting overall survival for patients with HCC. IDMS -integrated normograms provide a promising and quantitative tool for clinical risk management.Notably, a significant difference in microsatellite instability (MSI) was observed between the high- and low-risk groups. This indicates that patients in the high-risk group might have a better response to immunotherapy. Additionally, we predicted potential drugs targeting to these risk subgroups.

Conclusion: Our research developed an IDMS that could serve as an effective tool for patient stratification management and prognosis prediction. This signature could provide a reference for immunotherapy for patients with HCC and improve their prognosis.

Keywords: hepatocellular carcinoma, single-cell RNA-seq, machine learning, multi-omics, biomarkers, immunotherapy response

Introduction

The incidence of hepatocellular carcinoma (HCC) is on the rise, making it the third most prevalent cause of cancer-related mortality globally.[1,2] HCC typically presents with an insidious onset and rapid progression, contributing to a global five-year survival rate of less than 30%.[3] Over the past decades, despite the establishment of a multidisciplinary treatment approach—including surgery, radiotherapy, radiofrequency ablation, molecular targeted therapy, and transcatheter arterial chemoembolization—morbidity and mortality have remained largely unchanged, and the overall prognosis remains poor.[4] Given the significant heterogeneity of HCC, predictive, preventive, and personalized medicine (PPPM) represents a critical strategy for improving patient outcomes.[5] Although several biomarkers related to HCC have been

https://doi.org/10.2147/JHC.S574690
Received: 16 October 2025
Accepted: 28 February 2026
Published: 11 March 2026

© 2026 Wu et al. This work is published and licensed by Dove Medical Press Limited. The full terms of this license are available at https://www.dovepress.com/terms.php and incorporate the Creative Commons Attribution – Non Commercial (unported, v4.0) License (http://creativecommons.org/licenses/by-nc/4.0/). By accessing the work you hereby accept the Terms. Non-commercial uses of the work are permitted without any further permission from Dove Medical Press Limited, provided the work is properly attributed. For permission for commercial use of this work, please see paragraphs 4.2 and 5 of our Terms (https://www.dovepress.com/terms.php).

identified in recent years,[6–8] their clinical application remains limited. First, bulk transcriptomic analyses often fail to capture the pronounced heterogeneity of the tumor microenvironment (TME). Second, the robustness and generalizability of biomarker signatures developed across independent research groups show considerable variability in validation cohorts,[9,10] thereby diminishing their reliability for guiding individualized therapeutic decisions.

Currently, the intricate interactions between immune and cancer cells have emerged as a crucial focus of research efforts.[11,12] The TME consists of stromal cells, fibroblasts, immune cells, endothelial cells, and cancer cells.[13] The significant association of the immune cells and various tumor cells in the TME have been demonstrated.[14–16] Cancer cells can maintain an immunosuppressive microenvironment and evade antitumor responses through multiple mechanisms, such as secreting immunosuppressive factors or recruiting regulatory T cells (Tregs).[17,18] Meanwhile, chemotherapy, radiotherapy, and targeted therapies have been shown to enhance the immunogenicity of tumor cells and stimulate immune responses, thereby improving the efficacy of anticancer treatments.[19–21] These findings suggest that immune-related genes (IRGs) serve as pivotal molecular mediators bridging tumor heterogeneity and personalized therapeutic strategies, underscoring their substantial translational and research significance.

With the advancement of sequencing technology, the integrated analysis of multi-omics has provided a new approach for exploring the biological characteristics of tumors. Herein, we employed a multi-omics approach combined with machine learning to investigate the potential clinical value of IRGs in HCC and identify an immune-derived molecular signature (IDMS). Our research indicates that the IDMS could be used as an innovative biomarker for predicting the prognosis and immunotherapy response in HCC, thereby contributing to improving patient outcomes through targeted prevention and personalized medical care.

Materials and Methods

Acquisition and Processing of Data

The single-cell dataset GSE162616[22] was sourced from the Gene Expression Omnibus (GEO), converted into a Seurat object, and screened to retain high-quality cells (mitochondrial gene content < 20%, number of features > 300, $\log_{10}\text{-GenesPerUMI} > 0.80$). The functions "Harmony" was employed to eliminate the impact of batch effects, followed by data normalization. The "FindMarkers" function was then employed to identify single-cell differentially expressed genes (scDEGs) with |logFC| > 0.25 and an adjusted *p*-value < 0.05. Bulk RNA-seq data were downloaded from The Cancer Genome Atlas (TCGA), and the associated clinical information was retrieved from the UCSC Xena platform. After excluding samples lacking clinical information, a final cohort of 424 patients from The Cancer Genome Atlas Liver Hepatocellular Carcinoma (TCGA-LIHC) was obtained. To robustly identify key genes exhibiting consistent differential expression while minimizing the risk of overlooking biologically relevant molecules due to excessively stringent statistical thresholds, this study applied a DEG screening criterion of $|\log_2 FC| > 0.3$ and an adjusted p-value < 0.05. Additionally, three independent datasets (GSE14520, GSE76427, GSE121248) were obtained from GEO,[23–25] and a comprehensive dataset composed of 2,483 IRGs was retrieved from the ImmPort database. The full gene list is provided in Table 1.

Weighted Gene Co-Expression Network Analysis (WGCNA)

The "WGCNA"[26] was utilized to compute correlation coefficients between gene pairs and apply these weighted coefficients to construct a gene co-expression network consistent with a scale-free topology. Subsequently, the TCGA-LIHC dataset was analyzed to calculate gene variance, and the top 5,000 genes with the highest median absolute deviation were selected. We then assessed correlations between HCC and control groups across various modules. Genes within modules that exhibited an absolute correlation coefficient (|r|) greater than 0.50 were selected for further analysis.

Identification of Immune-Derived Molecule Signature and Prognosis Model

We obtained hub genes by intersecting of IRGs, scDEGs, DEGs and module genes from WGCNA. Hub genes were selected by five machine learning algorithms: the least absolute shrinkage and selection operator (LASSO), Boruta, random forest (RF), learning vector quantization (LVQ), and Bagged Trees. The overlapping genes detected by these five

Table 1 Immune-Related Genes (IRGs)

Immune-Related Genes				
AZGP1	APOBEC3H	CXCL11	SAA2	KIR3DL1
B2M	TMPRSS6	CXCL12	SBDS	KIR3DL2
CALR	SPINK5	CXCL13	SCG2	KIR2DL1
CANX	MARCO	CXCL14	SCGB3A1	KIR2DL2
CD1A	BECN1	CXCL16	SCT	KIR2DL3
CD1B	TNFSF11	CXCL17	AIMP1	KIR2DL4
CD1C	KNG1	CXCL2	SECTM1	KIR2DL5A
CD1D	CSK	CXCL3	SEMA3A	KLRC1
CD1E	KLRK1	CXCL5	SEMA3B	KLRC2
CD4	KCNH2	CXCL6	SEMA3C	KLRC3
CD8A	JUND	CXCL9	SEMA3D	KLRD1
CD8B	JAK1	CCN1	SEMA3E	PTPN6
CD74	CREB1	DEFA1	SEMA3F	PTPN11
CREB1	CLDN4	DEFA3	SEMA3G	ICAM1
CTSB	CCL28	DEFA5	SEMA4A	ICAM2
CTSE	RNASE3	DEFB1	SEMA4B	ITGAL
CTSL	RN7SL1	DEFB103B	SEMA4C	ITGB2
CTSS	IRF7	DEFB104A	SEMA4D	PTK2B
FCER1G	IREB2	DEFB4A	SEMA4F	VAV3
FCGRT	ILK	EDN1	SEMA4G	VAV1
PDIA3	IL18	EDN2	SEMA5A	VAV2
HFE	IL17A	EDN3	SEMA5B	RAC1
HLA-A	LTB4R	FGF10	SEMA6A	RAC2
HLA-B	APOBEC3A	FGF2	SEMA6B	RAC3
HLA-C	MASP2	HTN3	SEMA6C	PAK1
HLA-DMA	TRIM27	CXCL8	SEMA6D	MAP2K1
HLA-DMB	RELA	LECT2	SEMA7A	MAP2K2
HLA-DOA	IL7R	PF4	SLIT1	MAPK1
HLA-DOB	IL1A	PF4V1	SLIT2	MAPK3
HLA-DPA1	PTX3	PLAU	SLURP1	TNF
HLA-DPB1	IFNAR2	PPBP	SPP1	CSF2
HLA-DQA1	IFN1@	PPBPP1	SST	IFNG
HLA-DQA2	SYTL1	PROK2	STC1	KIR2DS1
HLA-DQB1	APOBEC3C	RNASE2	STC2	KIR2DS3
HLA-DRA	DDX17	SAA1	TAC1	KIR2DS4
HLA-DRB1	PTGS2	SAA2	TDGF1	KIR2DS5
HLA-DRB3	HTR1A	SBDS	TDGF1P3	NCR2
HLA-DRB4	SEPTIN7	SEMA3A	TG	TYROBP
HLA-DRB5	CD40LG	SEMA3B	TGFA	LCK
HLA-E	CD14	SEMA3C	TGFB1	FCGR3A
HLA-F	CD8A	SEMA3D	TGFB2	FCGR3B
HLA-G	CD4	SEMA3E	TGFB3	NCR1
HLA-H	MASP1	SEMA3F	THPO	NCR3
MR1	PROC	SEMA3G	TNC	FCER1G
HSPA1A	MAP2K2	SEMA4A	TNF	CD247
HSPA1B	MAP2K1	SEMA4B	TNFRSF11B	ZAP70
HSPA1L	HRG	SEMA4C	TNFSF10	SYK
HSPA2	NDRG1	SEMA4D	TNFSF11	LCP2
HSPA4	IRF9	SEMA4F	TNFSF12	LAT
HSPA5	TRIM22	SEMA4G	TNFSF13	PLCG1
HSPA6	LANCL1	SEMA5A	TNFSF13B	PLCG2

(*Continued*)

Table 1 (Continued).

Immune-Related Genes				
HSPA8	PPP4C	SEMA5B	TNFSF14	SH3BP2
HSP90AA1	HMOX1	SEMA6A	TNFSF15	PIK3CA
HSP90AB1	HMGB1	SEMA6B	TNFSF18	PIK3CB
ICAM1	HLA-B	SEMA6C	TNFSF4	PIK3CD
IFNA1	RNASE7	SEMA6D	TNFSF8	PIK3CG
IFNA2	ABCC4	SEMA7A	TNFSF9	PIK3R5
IFNA4	HGF	SLIT1	TOR2A	PIK3R1
IFNA5	HDAC1	SLIT2	TRH	PIK3R2
IFNA6	IFNLR1	TNC	TSHB	PIK3R3
IFNA7	PLSCR1	TYMP	TSLP	FYN
IFNA8	B2M	XCL1	TXLNA	SHC2
IFNA10	BACH2	XCL2	TYMP	SHC4
IFNA13	TANK	C5AR1	UCN	SHC3
IFNA14	PIK3CG	ACKR2	UCN2	SHC1
IFNA16	ARRB1	CCR1	UCN3	GRB2
IFNA17	RSAD2	CCR10	UTS2	SOS1
IFNA21	STAB2	CCR3	UTS2B	SOS2
IFNG	TBK1	CCR4	VEGFA	HRAS
KIR2DL1	PDYN	CCR5	VEGFB	KRAS
KIR2DL2	PDGFRB	CCR6	VEGFC	NRAS
KIR2DL3	PDCD1	CCR7	VGF	ARAF
KIR2DL4	PCSK2	CCR8	VIP	BRAF
KIR2DS1	PCSK1	CCR9	XCL1	RAF1
KIR2DS3	ARG2	ACKR4	XCL2	MICA
KIR2DS4	AQP9	CCRL2	ACVR1B	MICB
KIR2DS5	FASLG	CMKLR1	ACVR1C	ULBP3
KIR3DL1	APOH	CX3CR1	ACVR2A	ULBP2
KIR3DL2	BIRC5	CXCR3	ACVR2B	ULBP1
KLRC1	ANXA6	CXCR4	ACVRL1	KLRK1
KLRC2	IL22	CXCR5	ADCYAP1R1	HCST
KLRC3	VTN	CXCR6	ADIPOR1	CD48
KLRD1	VIM	ACKR3	ADIPOR2	CD244
LTA	VCAM1	CYSLTR1	ADRB1	PPP3CA
CIITA	PRDX1	CYSLTR2	ADRB2	PPP3CB
MICA	GFAP	ACKR1	AGTR1	PPP3CC
MICB	GBP2	EDNRA	AGTR2	CHP1
NFYA	ALB	EDNRB	AMHR2	PPP3R1
NFYB	SLC29A3	FPR1	ANGPT1	PPP3R2
NFYC	OAS1	FPR2	ANGPT4	CHP2
LGMN	AGER	FPR2	ANGPTL1	NFAT5
PSMB8	UNC93B1	GPR17	ANGPTL2	NFATC1
PSMC1	TNFSF4	GPR32	ANGPTL3	NFATC2
PSMC2	NOS1	GPR33	ANGPTL4	NFATC3
PSMC3	ACTG1	PTGDR2	ANGPTL6	NFATC4
PSMC4	ACTA1	C5AR2	APLNR	PRKCA
PSMC5	ACO1	CXCR1	AR	PRKCB
PSMC6	SERPINA3	CXCR2	AVPR1A	PRKCG
PSMD1	CXCR1	LTB4R	AVPR1B	SH2D1B
PSMD2	CCL15	LTB4R2	AVPR2	SH2D1A
PSMD3	CCL14	PLAUR	BMPR1A	IFNGR1

(*Continued*)

Table 1 (Continued).

Immune-Related Genes				
PSMD4	CCL4	PLXNA1	BMPR1B	IFNGR2
PSMD5	CCL16	PLXNA2	BMPR2	IFNA1
PSMD7	CCL19	PLXNA3	BRD8	IFNA2
PSMD8	CCL13	PLXNA4	C3AR1	IFNA4
PSMD10	CCL18	PLXNB1	C5AR1	IFNA5
PSMD11	CCL17	PLXNB2	CALCR	IFNA6
PSMD13	CCL26	PLXNB3	CALCRL	IFNA7
PSME1	CCL22	PLXNC1	ACKR2	IFNA8
PSME1	CCR3	PLXND1	CCR1	IFNA10
PSME2	CCL28	PTAFR	CCR10	IFNA13
PSME2	CCL4L1	ROBO1	CCR3	IFNA14
RELB	ACKR2	ROBO2	CCR4	IFNA16
RFX5	CCR7	ROBO3	CCR5	IFNA17
RFXAP	CCL27	RXFP3	CCR6	IFNA21
SLC10A2	CCR8	XCR1	CCR7	IFNB1
TAP1	ACKR4	ADIPOQ	CCR8	IFNAR1
TAP2	CCR10	ADM	CCR9	IFNAR2
TAPBP	CCL2	ADM2	ACKR4	TNFSF10
THBS1	CCL21	AGRP	CCRL2	TNFRSF10D
SEM1	CCL7	AGT	CD40	TNFRSF10C
KLRC4	CCL5	AMBN	CMKLR1	TNFRSF10B
AP3B1	CCL3	AMELX	CNTFR	TNFRSF10A
RFXANK	CCL20	AMH	CRHR1	FASLG
PSMD6	CCL11	ANGPTL5	CRHR2	FAS
PSME3	CCR5	ANGPTL7	CRIM1	GZMB
PSMD14	CCL23	APLN	CRLF1	PRF1
CLEC4M	CCL25	AREG	CRLF2	CASP3
IFI30	CCL1	MANF	CRLF3	BID
PROCR	CCL3L3	CDNF	CSF1R	CD3D
ADRM1	CCL4L2	ARTN	CSF2RA	CD3E
ECPAS	CXCL12	AVP	CSF2RB	CD3G
TRPC4AP	XCL1	AZU1	CSF3R	CD247
CD209	CCL8	BDNF	CX3CR1	CD4
UBXN1	CCL3L1	BMP1	CXCR3	CD8A
ERAP1	CCR1	BMP10	CXCR4	CD8B
TAPBPL	CCL24	BMP15	CXCR5	PTPRC
KIR2DL5A	XCL2	BMP2	CXCR6	LCK
ERAP2	CXCL1	BMP3	ACKR3	FYN
ULBP3	CXCL10	BMP4	CYSLTR1	ZAP70
ULBP2	CXCR4	BMP5	CYSLTR2	LCP2
ULBP1	CXCL2	BMP6	ACKR1	LAT
KIR3DL3	CXCR6	BMP7	EDNRA	ITK
RAET1E	CCR4	BMP8A	EDNRB	TEC
RAET1L	CXCL11	BMP8B	EGFR	NCK1
UBR1	TAFA5	BTC	ENG	NCK2
RAET1G	TAFA3	MYDGF	EPOR	VAV3
PDIA2	TAFA4	C3	ESR1	VAV1
HAMP	TAFA1	C5	ESR2	VAV2
PI3	TAFA2	CALCA	ESRRA	GRAP2
CAMP	CCL15-CCL14	CALCB	ESRRB	GRB2

(*Continued*)

Table 1 (Continued).

Immune-Related Genes				
DEFB4A	IL6	CAMP	ESRRG	PAK1
PPBP	TNF	CAT	FGFR1	PAK2
REG3G	IL1B	CCK	FGFR2	PAK3
CXCL14	IL18	CCL1	FGFR3	PAK4
CXCL16	PTK2B	CCL11	FGFR4	PAK6
SLPI	VEGFA	CCL13	FGFRL1	PAK5
CXCL8	IL4	CCL14	FLT1	RHOA
CXCL10	CDH1	CCL15-CCL14	FLT3	CDC42
CXCL9	CD40	CCL15	FLT4	PPP3CA
CXCL5	DEFB103B	CCL16	FPR1	PPP3CB
CXCL11	F2RL1	CCL17	FPR2	PPP3CC
CXCL6	MMP9	CCL18	FPR2	CHP1
CXCL1	LTBP1	CCL19	FSHR	PPP3R1
CXCL12	DEFB4A	CCL2	GALR2	PPP3R2
CXCL13	TNFSF10	CCL20	GALR3	CHP2
CXCL2	IL13	CCL21	GCGR	NFAT5
PF4	IL10	CCL22	GHR	NFATC1
XCL1	IL2	CCL23	GHRHR	NFATC2
CXCL3	PPARG	CCL24	GHSR	NFATC3
DEFB103B	FGR	CCL25	GIPR	NFATC4
CCL13	MIF	CCL26	GLP1R	SOS1
CCL1	CRP	CCL27	GLP2R	SOS2
DEFB1	JAK2	CCL28	GNRHR	HRAS
CCL8	IL1A	CCL3	GPER1	KRAS
ELANE	PTK2	CCL3L1	GPR17	NRAS
DEFB103A	PTGDR	CCL3P1	GPR32	FOS
DEFA3	CD86	CCL3L3	GPR33	JUN
DEFA1	HCK	CCL4	PTGDR2	CARD11
TMSB10	ARRB1	CCL4L2	C5AR2	BCL10
DEFA6	GNAI1	CCL4L1	HNF4A	MALT1
DEFA5	VDR	CCL5	HNF4G	CHUK
DEFA4	OLR1	CCL7	HTR3A	IKBKB
LCN2	GRK2	CCL8	HTR3B	IKBKG
LCN1	TXK	CD320	HTR3C	NFKB1
COLEC10	RNASE2	CD40LG	HTR3D	RELA
BPI	CD79A	CD70	HTR3E	NFKBIA
S100A9	CD79B	ADA2	IFNAR1	NFKBIB
S100A8	LYN	CER1	IFNAR2	NFKBIE
DCD	SYK	CGA	IFNGR1	CD28
LCN6	BTK	CGB3	IFNGR2	ICOS
S100A12	BLNK	CGB1	IGF1R	CD40LG
HTN3	VAV3	CGB2	IGF2R	PIK3R5
LCN8	VAV1	CGB5	IL10RA	PIK3R1
DEFA1B	VAV2	CGB7	IL10RB	PIK3R2
CCR10	RAC1	CGB8	IL11RA	PIK3R3
CELA1	RAC2	CHGA	IL12RB1	PIK3CA
DEFB106A	RAC3	CHGB	IL12RB2	PIK3CB
PENK	PPP3CA	CKLF	IL13RA1	PIK3CD
BPIFC	PPP3CB	CLCF1	IL13RA2	PIK3CG
MMP12	PPP3CC	CLEC11A	IL15RA	AKT3

(*Continued*)

Table 1 (Continued).

Immune-Related Genes				
BPIFB6	CHP1	CMA1	IL2RB	AKT1
LEAP2	PPP3R1	CMTM1	IL17RA	AKT2
SFTPD	PPP3R2	CMTM2	IL17RB	MAP3K8
LCN9	CHP2	CMTM3	IL17RC	MAP3K14
BPIFB2	NFAT5	CMTM4	IL17RD	PDCD1
PTGDS	NFATC1	CMTM5	IL17RE	CTLA4
TMSB4X	NFATC2	CMTM6	IL18R1	PTPN6
PGLYRP1	NFATC3	CMTM7	IL18RAP	CBLC
ZC3HAV1	NFATC4	CMTM8	IL1R1	CBL
TMSB15A	HRAS	CNTF	IL1R2	CBLB
S100B	KRAS	CORT	IL1RAP	IL2
S100A13	NRAS	CRH	IL1RL1	IL4
S100A6	FOS	CSF1	IL1RL2	IL5
DEFB119	JUN	CSF2	IL20RA	IL10
DEFB107A	CARD11	CSF3	IL20RB	IFNG
DEFB105A	BCL10	CSH1	IL21R	CSF2
SERPIND1	MALT1	CSH2	IL22RA1	TNF
DEFB129	CHUK	CSHL1	IL22RA2	CDK4
DEFB127	IKBKB	CSPG5	IL23R	RASGRP1
S100P	IKBKG	CTF1	IL27RA	PDK1
S100A7	NFKB1	CCN2	IFNLR1	PLCG1
DEFB104A	RELA	CTSG	IL2RA	PRKCQ
DEFB126	NFKBIA	CX3CL1	IL2RB	TRAC
DEFB106B	NFKBIB	CXCL1	IL2RG	TRAJ1
DEFB104B	NFKBIE	CXCL10	IL31RA	TRAJ2
DEFB107B	CD81	CXCL11	IL3RA	TRAJ3
PGLYRP3	CD19	CXCL12	IL4R	TRAJ4
PGLYRP2	CR2	CXCL13	IL5RA	TRAJ5
S100A10	PIK3R5	CXCL14	IL6R	TRAJ6
S100A2	PIK3R1	CXCL16	IL7R	TRAJ7
DEFB125	PIK3R2	CXCL17	CXCR1	TRAJ8
DEFB123	PIK3R3	CXCL2	CXCR2	TRAJ9
DEFB105B	PIK3CA	CXCL3	IL9R	TRAJ10
DEFB132	PIK3CB	CXCL5	INSR	TRAJ11
BPIFB3	PIK3CD	CXCL6	KDR	TRAJ12
LCN12	PIK3CG	CXCL9	LEPR	TRAJ13
PGLYRP4	AKT3	CCN1	LGR4	TRAJ14
S100A11	AKT1	DEFA1	LGR5	TRAJ15
S100A5	AKT2	DEFA3	LGR6	TRAJ16
S100A3	GSK3B	DEFA5	LHCGR	TRAJ17
S100A1	INPP5D	DEFB1	LIFR	TRAJ18
DEFB128	CD22	DEFB103B	LTB4R	TRAJ19
DEFB108B	CD72	DEFB104A	LTB4R2	TRAJ20
HTN1	PTPN6	DEFB4A	LTBR	TRAJ21
LMBR1L	LILRB3	DKK1	MC1R	TRAJ22
S100A7A	FCGR2B	EBI3	MC2R	TRAJ23
DEFB118	RASGRP3	EDN1	MC3R	TRAJ24
COLEC12	PLCG2	EDN2	MC4R	TRAJ25
TMSB4Y	PRKCB	EDN3	MCHR1	TRAJ26
DEFB131A	IFITM1	EGF	MCHR2	TRAJ27

(*Continued*)

Table 1 (Continued).

Immune-Related Genes				
DEFB134	IGH	EPGN	MET	TRAJ28
DEFB130A	IGHA1	EPO	MLNR	TRAJ29
DEFB124	IGHA2	EREG	MPL	TRAJ30
DEFB121	IGHD	ESM1	MTNR1A	TRAJ31
DEFB116	IGHD1-1	FAM3B	MTNR1B	TRAJ32
DEFB115	IGHD1-14	FAM3C	NGFR	TRAJ33
DEFB114	IGHD1-20	FAM3D	NMBR	TRAJ34
DEFB113	IGHD1-26	FASLG	NPR1	TRAJ35
DEFB112	IGHD1-7	FGF1	NPR3	TRAJ36
DEFB110	IGHD2-15	FGF10	NR0B1	TRAJ37
TMSB15B	IGHD2-2	FGF11	NR0B2	TRAJ38
DEFB133	IGHD2-21	FGF12	NR1D1	TRAJ39
S100Z	IGHD2-8	FGF13	NR1D2	TRAJ40
MAVS	IGHD3-10	FGF14	NR1H2	TRAJ41
TMSB4XP8	IGHD3-16	FGF16	NR1H3	TRAJ42
S100A14	IGHD3-22	FGF17	NR1H4	TRAJ43
LCN10	IGHD3-3	FGF18	NR1I2	TRAJ44
S100A16	IGHD3-9	FGF19	NR1I3	TRAJ45
DEFB136	IGHD4-11	FGF2	NR2C1	TRAJ46
DEFB135	IGHD4-17	FGF20	NR2C2	TRAJ47
DEFB117	IGHD4-23	FGF21	NR2E1	TRAJ48
DEFB110	IGHD4-4	FGF22	NR2E3	TRAJ49
ZC3HAV1L	IGHD5-12	FGF23	NR2F1	TRAJ50
S100A7L2	IGHD5-18	FGF3	NR2F2	TRAJ52
MBL3P	IGHD5-24	FGF4	NR2F6	TRAJ53
DEFB4B	IGHD5-5	FGF5	NR3C1	TRAJ54
BPIFB4	IGHD6-13	FGF6	NR3C2	TRAJ56
IFNAR1	IGHD6-19	FGF7	NR4A1	TRAJ57
AZU1	IGHD6-25	FGF8	NR4A2	TRAJ58
DEFB131B	IGHD6-6	FGF9	NR4A3	TRAJ59
DEFA1A3	IGHD7-27	VEGFD	NR5A1	TRAJ61
LCN1P1	IGHE	FIGNL2	NR5A2	TRAV1-1
S100G	IGHG1	FLT3LG	NR6A1	TRAV1-2
DEFA7P	IGHG2	FSHB	NRP1	TRAV2
DEFB130B	IGHG3	GAL	NRP2	TRAV3
DEFB108F	IGHG4	GALP	OGFR	TRAV4
DEFB131C	IGHJ1	GAST	OPRD1	TRAV5
TCHHL1	IGHJ2	GCG	OPRK1	TRAV7
TINAGL1	IGHJ3	GDF1	OPRL1	TRAV8-1
IFNGR1	IGHJ4	GDF10	OPRM1	TRAV8-2
SLC22A17	IGHJ5	GDF11	OSMR	TRAV8-3
WFIKKN1	IGHJ6	GDF15	OXTR	TRAV8-4
WFDC2	IGHM	GDF2	PGR	TRAV8-6
IL6	IGH	GDF3	PGRMC2	TRAV8-7
UMODL1	IGHV1-18	GDF5	PLAUR	TRAV9-1
TGFB1	IGHV1-2	GDF6	PLXNA1	TRAV9-2
PF4V1	IGHV1-24	GDF7	PLXNA2	TRAV10
MMP9	IGHV1-3	GDF9	PLXNA3	TRAV12-1
ANOS1	IGHV1-45	GDNF	PLXNA4	TRAV12-2
TLR4	IGHV1-46	GH1	PLXNB1	TRAV12-3

(*Continued*)

Table 1 (Continued).

Immune-Related Genes				
IFNG	IGHV1-58	GH2	PLXNB2	TRAV13-1
SPAG11B	IGHV1-69	GHRH	PLXNB3	TRAV13-2
A2M	IGHV1-8	GHRL	PLXNC1	TRAV14DV4
CTSL	IGHV1-38-4	GIP	PLXND1	TRAV16
NFKB1	IGHV1-69-2	GKN1	PPARA	TRAV17
APOBEC3G	IGHV2-26	GMFB	PPARD	TRAV18
FABP6	IGHV2-5	GMFG	PPARG	TRAV19
NOD2	IGHV2-70	GNRH1	PRLHR	TRAV20
MBL2	IGHV3-11	GNRH2	PRLR	TRAV21
SFTPA1	IGHV3-13	GPHA2	PTAFR	TRAV22
RBP1	IGHV3-15	GPHB5	PTGDR	TRAV23DV6
TLR2	IGHV3-16	GPI	PTGDS	TRAV24
SLC40A1	IGHV3-20	GREM1	PTGER1	TRAV25
PLAU	IGHV3-21	GREM2	PTGER2	TRAV26-1
IL1B	IGHV3-23	GRN	PTGER3	TRAV26-2
PAEP	IGHV3-30	GRP	PTGER4	TRAV27
HJV	IGHV3-30-3	GUCA2A	PTGFR	TRAV29DV5
MUC5AC	IGHV3-30-5	HAMP	PTH1R	TRAV30
CTSS	IGHV3-33	HBEGF	PTH2R	TRAV34
OBP2A	IGHV3-35	HDGF	RARA	TRAV35
PLTP	IGHV3-38	HDGFL3	RARB	TRAV36DV7
MX1	IGHV3-43	HGF	RARG	TRAV38-1
DDX58	IGHV3-48	HTN3	ROBO1	TRAV38-2DV8
IFNL1	IGHV3-49	IAPP	ROBO2	TRAV39
IRF3	IGHV3-53	IFNA1	ROBO3	TRAV40
SFTPA2	IGHV3-64	IFNA10	RORA	TRAV41
LPA	IGHV3-66	IFNA13	RORB	TRBC1
LBP	IGHV3-7	IFNA14	RORC	TRBC2
RBP4	IGHV3-72	IFNA16	RXFP1	TRBD1
SFTPA1	IGHV3-73	IFNA17	RXFP2	TRBD2
NOX4	IGHV3-74	IFNA2	RXFP3	TRBJ1-1
LTF	IGHV3-9	IFNA21	RXRA	TRBJ1-2
IFNB1	IGHV3-38-3	IFNA4	RXRB	TRBJ1-3
RBP5	IGHV3-69-1	IFNA5	RXRG	TRBJ1-4
FABP7	IGHV4-28	IFNA6	S1PR1	TRBJ1-5
FABP5	IGHV4-30-1	IFNA7	S1PR2	TRBJ1-6
FABP3	IGHV4-30-2	IFNA8	SCTR	TRBJ2-1
FABP2	IGHV4-30-4	IFNB1	SDC1	TRBJ2-2
FABP4	IGHV4-31	IFNE	SDC2	TRBJ2-3
R3HDML	IGHV4-34	IFNG	SDC3	TRBJ2-4
BPIFA3	IGHV4-39	IFNK	SDC4	TRBJ2-5
BPIFB1	IGHV4-4	IFNW1	SORT1	TRBJ2-6
OASL	IGHV4-59	IGF1	SSTR1	TRBJ2-7
CRABP2	IGHV4-61	IGF2	SSTR2	TRBV2
CRABP1	IGHV4-38-2	IL10	SSTR5	TRBV3-1
RBP7	IGHV5-51	IL11	ST2	TRBV4-1
DUOX1	IGHV5-10-1	IL12A	TACR1	TRBV4-2
OBP2B	IGHV6-1	IL12B	TEK	TRBV4-3
RBP2	IGHV7-4-1	IL13	TGFBR1	TRBV5-1
LCN15	IGHV7-81	IL15	TGFBR2	TRBV5-4

(*Continued*)

Table 1 (Continued).

Immune-Related Genes				
CETP	IGK	IL16	TGFBR3	TRBV5-5
FABP12	IGKC	IL17A	THRA	TRBV5-6
FABP9	IGKDEL	IL17B	THRB	TRBV5-7
BPIFA1	IGKJ	IL17C	TIE1	TRBV5-8
LCNL1	IGKJ1	IL17D	TNFRSF10A	TRBV6-1
C8G	IGKJ2	IL17F	TNFRSF10B	TRBV6-2
SPAG11A	IGKJ3	IL18	TNFRSF10C	TRBV6-3
PI15	IGKJ4	IL19	TNFRSF10D	TRBV6-4
NOX1	IGKJ5	IL1A	TNFRSF11A	TRBV6-5
PMP2	IGKV@	IL1B	TNFRSF12A	TRBV6-6
APOD	IGKV1-12	IL1F10	TNFRSF13B	TRBV6-7
ORM2	IGKV1-13	IL36RN	TNFRSF13C	TRBV6-8
ORM1	IGKV1-16	IL36A	TNFRSF14	TRBV6-9
TNF	IGKV1-17	IL37	TNFRSF17	TRBV7-2
CTSG	IGKV1-27	IL36B	TNFRSF18	TRBV7-3
PRTN3	IGKV1-33	IL36G	TNFRSF19	TRBV7-4
MAPK1	IGKV1-37	IL1RN	TNFRSF1A	TRBV7-6
PML	IGKV1-39	IL2	TNFRSF1B	TRBV7-7
AEN	IGKV1-5	IL20	TNFRSF21	TRBV7-8
CYBB	IGKV1-6	IL21	TNFRSF25	TRBV7-9
BPIFA2	IGKV1-8	IL22	TNFRSF4	TRBV9
ISG20	IGKV1-9	IL23A	TNFRSF6B	TRBV10-1
BCL3	IGKV1D-12	IL24	TNFRSF8	TRBV10-2
ISG20L2	IGKV1D-13	IL25	TNFRSF9	TRBV10-3
NOX5	IGKV1D-16	IL26	TRHR	TRBV11-1
NOX3	IGKV1D-17	IL27	TSHR	TRBV11-2
DUOX2	IGKV1D-33	IFNL2	TUBB3	TRBV11-3
TLR3	IGKV1D-37	IFNL3	VDR	TRBV12-3
TFRC	IGKV1D-39	IFNL1	VIPR1	TRBV12-4
IFIH1	IGKV1D-42	IL3	VIPR2	TRBV12-5
LRP1	IGKV1D-43	IL31	XCR1	TRBV13
TRIM5	IGKV1D-8	IL32	IFNA10	TRBV14
IDO1	IGKV2-24	IL33	IFNA13	TRBV15
GDF15	IGKV2-28	IL34	IFNA14	TRBV16
NEDD4	IGKV2-30	IL4	IFNA16	TRBV17
ADIPOQ	IGKV2-40	IL5	IFNA17	TRBV18
STAT3	IGKV2D-24	IL6	IFNA2	TRBV19
STAT1	IGKV2D-28	IL6ST	IFNA21	TRBV20-1
IFNL2	IGKV2D-29	IL7	IFNA4	TRBV24-1
SOCS3	IGKV2D-30	CXCL8	IFNA5	TRBV25-1
SEMG1	IGKV2D-40	IL9	IFNA6	TRBV27
TNFSF10	IGKV3-11	INHA	IFNA7	TRBV28
CCL20	IGKV3-15	INHBA	IFNA8	TRBV29-1
SOCS1	IGKV3-20	INHBB	IFNB1	TRBV30
RNASEL	IGKV3-7	INHBC	IFNE	TRDC
IRF1	IGKV3D-11	INHBE	IFNG	TRDD1
IL15	IGKV3D-15	INS	IFNK	TRDD2
APOBEC3F	IGKV3D-20	INS-IGF2	IFNW1	TRDD3
PLAAT4	IGKV3D-7	INSL3	IFNAR2	TRDJ1
CHIT1	IGKV4-1	INSL4	IFNGR1	TRDJ2

(*Continued*)

Table 1 (Continued).

Immune-Related Genes				
IFNA1	IGKV5-2	INSL5	IFNGR2	TRDJ3
CD40	IGKV6-21	INSL6	IL11	TRDJ4
TLR7	IGKV6D-21	JAG1	IL12A	TRDV1
PPIA	IGKV6D-41	JAG2	IL12B	TRDV2
HFE	IGL	FGF7P6	IL13	TRDV3
ZYX	IGLC1	FGF7P3	IL15	TRGV9
NLRX1	IGLC2	KITLG	IL16	TRGV8
PGC	IGLC3	KL	IL17A	TRGV5
VEGFA	IGLC6	LACRT	IL17B	TRGV4
IKBKE	IGLC7	LECT2	IL17C	TRGV3
ISG15	IGLJ	LEFTY1	IL17D	TRGV2
DHX58	IGLJ1	LEFTY2	IL17F	TRGJP2
TNFAIP3	IGLJ2	LEP	IL18	TRGJP1
TFR2	IGLJ3	LHB	IL19	TRGJP
FCN2	IGLJ4	LIF	IL1A	TRGJ2
MUC4	IGLJ5	LRSAM1	IL1B	TRGJ1
F2R	IGLJ6	LTA	IL1F10	TRGC2
ELN	IGLJ7	LTB	IL36RN	TRGC1
IL27	IGLV@	LTBP1	IL36A	TRAV6
MAPT	IGLV1-36	LTBP2	IL37	BMP1
LYZ	IGLV1-40	LTBP3	IL36B	BMP10
CCL5	IGLV1-44	LTBP4	IL36G	BMP15
LEP	IGLV1-47	MDK	IL1RN	BMP2
CYLD	IGLV1-50	MIA	IL2	BMP3
KLKB1	IGLV1-51	MIF	IL20	BMP4
CST4	IGLV10-54	MLN	IL21	BMP5
CSRP1	IGLV11-55	MSTN	IL22	BMP6
MAPK14	IGLV2-11	NAMPT	IL23A	BMP7
JUN	IGLV2-14	NDP	IL24	BMP8A
ITGAV	IGLV2-18	NENF	IL25	BMP8B
IRF5	IGLV2-23	NGF	IL26	GDF1
CCR6	IGLV2-33	NMB	IL27	GDF10
IL12B	IGLV2-8	NODAL	IFNL2	GDF11
TLR8	IGLV3-1	CCN3	IFNL3	GDF15
GNLY	IGLV3-10	NPFF	IFNL1	GDF2
CD81	IGLV3-12	NPPA	IL3	GDF3
EIF2AK2	IGLV3-16	NPPB	IL31	GDF5
APOM	IGLV3-19	NPPC	IL32	GDF6
CACYBP	IGLV3-21	NPY	IL33	GDF7
NOD1	IGLV3-22	NRG1	IL34	GDF9
MAPK8	IGLV3-25	NRG2	IL4	GDNF
MAPK3	IGLV3-27	NRG3	IL5	INHA
BST2	IGLV3-32	NRG4	IL6	INHBA
BPHL	IGLV3-9	NRTN	IL6ST	INHBB
PLA2G2A	IGLV4-3	NTF3	IL7	INHBC
GRN	IGLV4-60	NTF4	CXCL8	INHBE
NEWENTRY	IGLV4-69	NTS	IL9	LEFTY1
PDGFRA	IGLV5-37	NUDT6	TXLNA	LEFTY2
GNAI1	IGLV5-39	OGN	IL10RA	NODAL
WNT5A	IGLV5-45	OSGIN1	IL10RB	TGFB1

(*Continued*)

Table 1 (Continued).

Immune-Related Genes				
FURIN	IGLV5-48	OSM	IL11RA	TGFB2
ADAR	IGLV5-52	OSTN	IL12RB1	TGFB3
TYK2	IGLV6-57	OXT	IL12RB2	ACVR1B
NOS2	IGLV7-43	ENDOU	IL13RA1	ACVR1C
TRAF3	IGLV7-46	PDGFA	IL13RA2	ACVR2A
TPT1	IGLV8-61	PDGFB	IL15RA	ACVR2B
TPM2	IGLV9-49	PDGFC	IL2RB	ACVRL1
NEO1	C3	PDGFD	IL17RA	AMHR2
AHNAK	C5	PDGFRA	IL17RB	BMPR1A
TLR1	CAMP	PDGFRB	IL17RC	BMPR1B
TK2	CCL1	PDGFRL	IL17RD	BMPR2
PRDX2	CCL11	PDYN	IL17RE	TGFBR1
MX2	CCL13	PENK	IL18R1	TGFBR2
FGF2	CCL14	PF4	IL18RAP	TGFBR3
FGA	CCL15-CCL14	PF4V1	IL1R1	TNFRSF11B
TCF7L2	CCL15	PGF	IL1R2	TNFSF10
F2RL1	CCL16	PLAU	IL1RAP	TNFSF11
TKFC	CCL17	PMCH	IL1RL1	TNFSF12
MSR1	CCL18	PNOC	IL1RL2	TNFSF13
NFKBIZ	CCL19	POMC	IL20RA	TNFSF13B
LMBR1	CCL2	PPBP	IL20RB	TNFSF14
EPPIN	CCL20	PPBPP1	IL21R	TNFSF15
SRC	CCL21	PPBPP2	IL22RA1	TNFSF18
MPO	CCL22	PPY	IL22RA2	TNFSF4
ELAVL1	CCL23	PRL	IL23R	TNFSF8
ROBO3	CCL24	PRLH	IL27RA	TNFSF9
SP1	CCL25	PROK1	IFNLR1	TNFRSF10B
SOD1	CCL26	PROK2	IL2RA	TNFRSF10C
PDF	CCL27	PSPN	IL2RB	TNFRSF10D
DLL4	CCL28	PTH	IL2RG	TNFRSF11A
ECD	CCL3	PTH2	IL31RA	TNFRSF12A
SLC11A1	CCL3L1	PTHLH	IL3RA	TNFRSF13B
DMBT1	CCL3P1	PTN	IL4R	TNFRSF13C
STING1	CCL3L3	PYY	IL5RA	TNFRSF14
SKIV2L	CCL4	QRFP	IL6R	TNFRSF17
SEMG2	CCL4L2	RABEP1	IL7R	TNFRSF18
LTA	CCL4L1	RABEP2	CXCR1	TNFRSF19
DES	CCL5	REG1A	CXCR2	TNFRSF1A
DCK	CCL7	RETN	IL9R	TNFRSF1B
DAXX	CCL8	RETNLB	ST2	TNFRSF21
TNFRSF10A	CKLF	RLN1	HLA-A	TNFRSF25
TNFRSF10B	CMA1	RLN2	HLA-B	TNFRSF4
EED	CTSG	RLN3	HLA-C	TNFRSF6B
CCL4	CX3CL1	RNASE2	HLA-E	TNFRSF8
LIMS1	CXCL1	S100A6	HLA-G	TNFRSF9
LALBA	CXCL10	SAA1		

algorithms were utilized to determine the IDMS for HCC. Subsequently, using GSE14520 as the validation set, six machine learning algorithms—LogitBoost, support vector machine (SVM), Naive Bayes (NB), RF, k-nearest neighbors (KNN), and AdaBoost—were applied to construct HCC classification models using the TCGA-LIHC dataset. Five-fold cross-validation was performed on each model to optimize para-meters. To guarantee the reliability of the results, the

optimization process was repeated ten times, each time using different random seeds for the resampling. The model that yielded the highest area under the curve (AUC) on the GSE14520 validation set was selected as the optimal prognosis model. In addition, for assessing the predictive performance of the prognosis model, we compared the predictive capabilities of IDMS with those of previously reported HCC markers across the TCGA-LIHC, GSE76427, GSE121248, and GSE14520.

Survival Analysis and Establishing a Nomogram

To evaluate the prognostic efficacy of the model, the TCGA-LIHC dataset was stratified into high- and low-risk groups based on the median value of predicted values. Kaplan-Meier survival curves were identify differences in OS between the two groups.Time-dependent receiver operating characteristic (ROC) curves were generated to calculate AUC using "survivalROC". Furthermore, clinical characteristics and the predicted value of the diagnostic model were incorporated into multivariate Cox regression analysis. Based on these factors, a nomogram was developed to assess the risk of HCC, and the predictive performance of the nomogram was evaluated using calibration curves.

Gene Set Enrichment Analysis (GSEA) and Gene Set Variation Analysis (GSVA)

The "limma" and "clusterProfiler" packages were used to perform GSEA and GSVA on all genes in the TCGA-LIHC dataset. Functional enrichment differences between the high- and low-risk groups were then calculated. Genes with an adjusted p-value < 0.05 were considered statistically significant. All p-value adjustments were performed using the Benjamini-Hochberg method.

Analysis of Mutation and Immune Characteristics

SM data for HCC patients in the TCGA-LIHC dataset were processed using "VarScan". CNV analysis was performed using GISTIC2.0.[27] The Mann–Whitney U-test was used to evaluate differences in microsatellite instability (MSI), tumor mutation burden (TMB), and tumor immune dysfunction and exclusion (TIDE) scores among the subgroups.

Analysis of Immune Infiltration and Drug Sensitivity

The abundance of immune cell types was evaluated using CIBERSORT. The Drug–Gene Interaction Database (DGIdb) was used to explore interactions between IDMS and small-molecule compounds for identifying potential therapeutic candidates. Predicted binding sites from CB-Dock2 (https://cadd.labshare.cn/cb-dock2/php/index.php) were used to visualize molecular docking between candidate drugs and their target proteins.

Quantitative Real-time Polymerase Chain Reaction (qRT-PCR)

Total RNA was isolated using the RNAiso plus RNA extraction kit (Takara,9019, Japan), following the manufacturer's guidelines.The PrimeScript RT reagent kit (Perfect Real Time, Takara, RR037A, Japan) was employed to conduct reverse transcription. qRT-PCR was conducted using the LineGene 9600Plus fluorescence qPCR system (Bioer, Hangzhou, China). Cycle threshold (Ct) values for each gene were normalized using GAPDH as an internal reference, and relative gene expression was calculated using the comparative Ct method. Data were derived from three biological replicates and analyzed employing Student's t-test. The primer sequences and detailed cell lines are available in Tables 2 and 3.

Statistical Analysis

All data analyses were carried out with R software version 4.2.1 and the corresponding R packages. Student's t-test was employed to compare data that followed a normal distribution,while the Mann–Whitney U-test was used for data with non-normal distribution.

Table 2 The Source of Cell Lines

Cell Line	Source
THLE-2	CL-0833,Pricella Biotechnology Co., Ltd, Wuhan,China,2024
MHCC97H	SCSP-5092,Chinese Academy of Sciences,Beijing,China,2024
HCCLM3	BNCC342335,BeNa Culture Collection,Beijing,China,2024
Hep3B	BNCC360312,BeNa Culture Collection,Beijing,China,2024
BEL-7405	BNCC359897,BeNa Culture Collection,Beijing,China,2024

Table 3 Primer Sequences for IDMS

HSP90AA1 FW (H)	AACCTATGGGTCGTGGAACA
HSP90AA1 RV (H)	CAGCCTCATCATCGCTTACTT
NR4A1 FW (H)	CACGGACCAGGGAAGTGAGGAGA
NR4A1 RV (H)	GGACAAGAGGCGGCGAAACC
S100A10 FW (H)	CTTATCAGGGAGGAGCGAACT
S100A10 RV (H)	AAAGGCTACTTAACAAAGGAGGA
GAPDH FW (H)	ATGACATCAAGAAGGTGGTGAAGCAGG
GAPDH RV (H)	GCGTCAAAGGTGGAGGAGTGGGT
TMSB10 FW (H)	AAGAACACCCTGCCGACCAAA
TMSB10 RV (H)	GGCTCTTCCTCCACATCACGACT

Results

Research Process and Immune-Derived Genes in Single-Cell Transcriptome

The overall analysis workflow is shown in Figure 1A. Following quality control of scRNA-seq dataset, a total of 61,763 cells were retained. After being annotated with the "SingleR" package, we identified six distinct cell types and 590 scDEGs (Figure 1B and C).

Identification of the Hub Genes in Bulk RNA-Seq

Using the "limma" R package, we identified 10,533 DEGs in the TCGA-LIHC dataset, including 9,175 upregulated and 1,358 downregulated genes (Figure 2A). Co-expression modules were constructed using WGCNA, and the optimal soft-thresholding power was determined to be 8 (Figure 2B). Correlations between module eigengenes and the HCC and control groups were evaluated based on the expression profiles of the top 5,000 genes with the highest median absolute deviation (MAD) (Figure 2C–E). And three modules containing genes were filtrated, namely: MEblue, MEblack, and MEbrown. Subsequently, seven hub genes—NR4A2, TMSB10, S100A10, NR4A1,MAP3K8, HSPA1A, and HSP90AA1—were identified for further investigation (Figure 2F).

Screening Immune-Derived Molecular Signature by Multi-Machine Learning

Five machine learning algorithms were applied to select the biological significance biomarkers in hub genes (Table 4). Seven genes (HSP90AA1, HSPA1A, MAP3K8, NR4A1, NR4A2, S100A10, TMSB10) were identified using LASSO algorithm (Figure 3A). Both the Boruta and LVQ algorithms selected seven genes, respectively (Figure 3B and C).

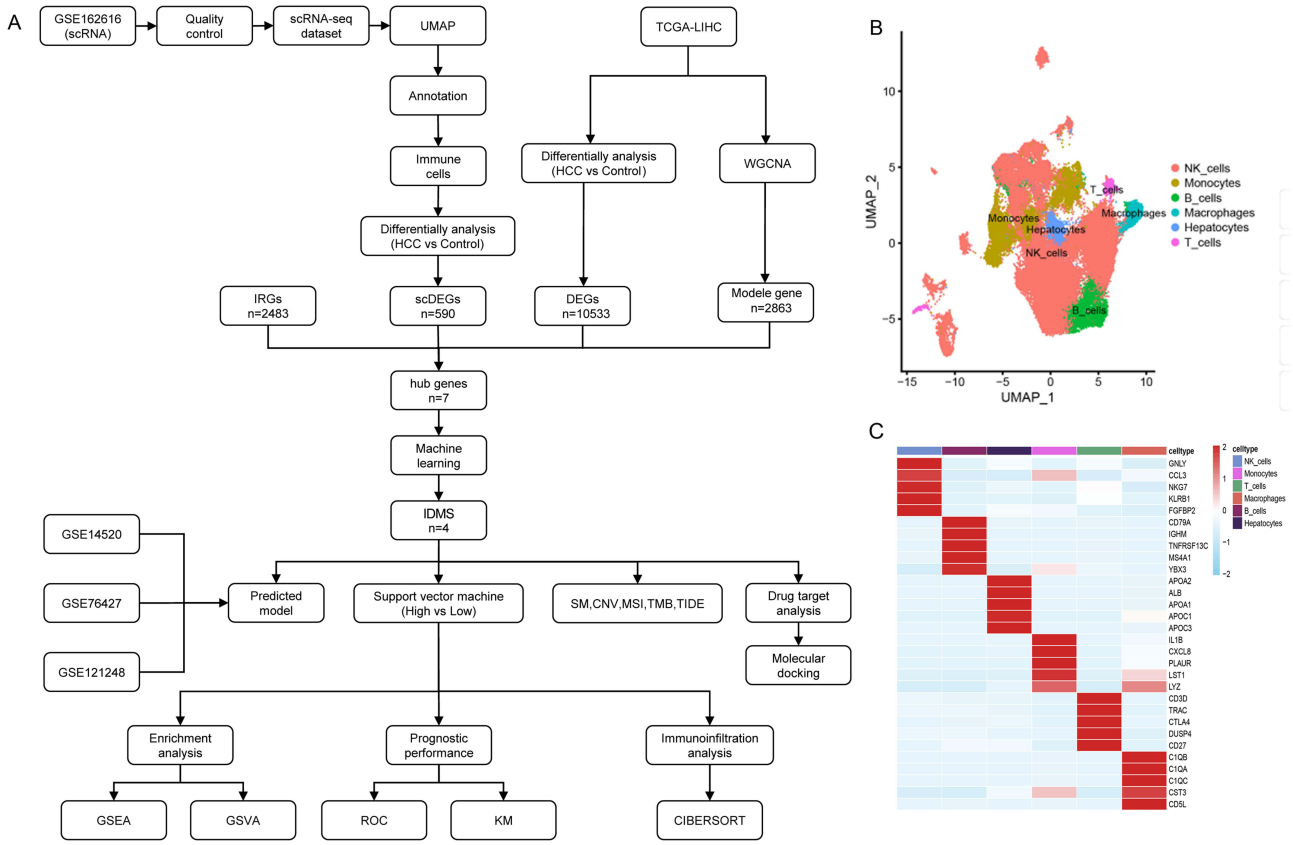

Figure 1 Flowchart and clustering analysis of scRNA-seq (**A**) Overall analytical workflow. (**B**) UMAP plot illustrating six identified cell types. (**C**) Heatmap of marker genes for each cell type.

Bagged Trees identified four genes (S100A10, NR4A1, HSP90AA1, TMSB10) (Figure 3D), while the random forest algorithm identified six genes (HSP90AA1, S100A10, NR4A1, NR4A2, TMSB10, MAP3K8) (Figure 3E). Finally, we selected four overlapping genes as IDMS: HSP90AA1, NR4A1, S100A10, and TMSB10 (Figure 3F).

Identifying Prognosis Model by Integrated Machine Learning Analysis

Based on 4 characteristic genes identified above, we evaluated the predict potential of IDMS using 6 machine learning algorithms. The ROC curve analysis showed that SVM model could effectively distinguish HCC from the control groups, with an AUC value of 0.957 (Figure 4A and B). Furthermore, compared with four previously published HCC signatures developed by Liu et al, Shen et al, Qian et al, and Zhou et al,[22,28–30] the IDMS exhibited consistently high predictive accuracy across multiple independent datasets, including TCGA-LIHC, GSE76427, and GSE14520. In the GSE121248 dataset, the predictive performance of IDMS was comparable to that of the biomarkers developed by Qian et al, with a higher meanAUC observed for IDMS (Figure 4C and D). Based on these findings, we selected the SVM as the optimal risk model for HCC. Subsequently, we performed annotation of IDMS and found that these genes were located on human chromosomes 1, 2, 12, and 14 (Figure 4E).

Predictive Performance of SVM for HCC Prognosis

Using the median predicted value generated by the SVM model, a risk factor plot was created to visualize stratification between high- and low-risk groups (Figure 5A). The analysis showed that the expression levels of S100A10, TMSB10 and HSP90AA1 are significantly elevated in the high-risk group, and the mortality is also higher (Figure 5B). Kaplan-Meier survival analysis revealed that increased expression of these genes, along with the predicted value, was

https://doi.org/10.2147/JHC.S574690

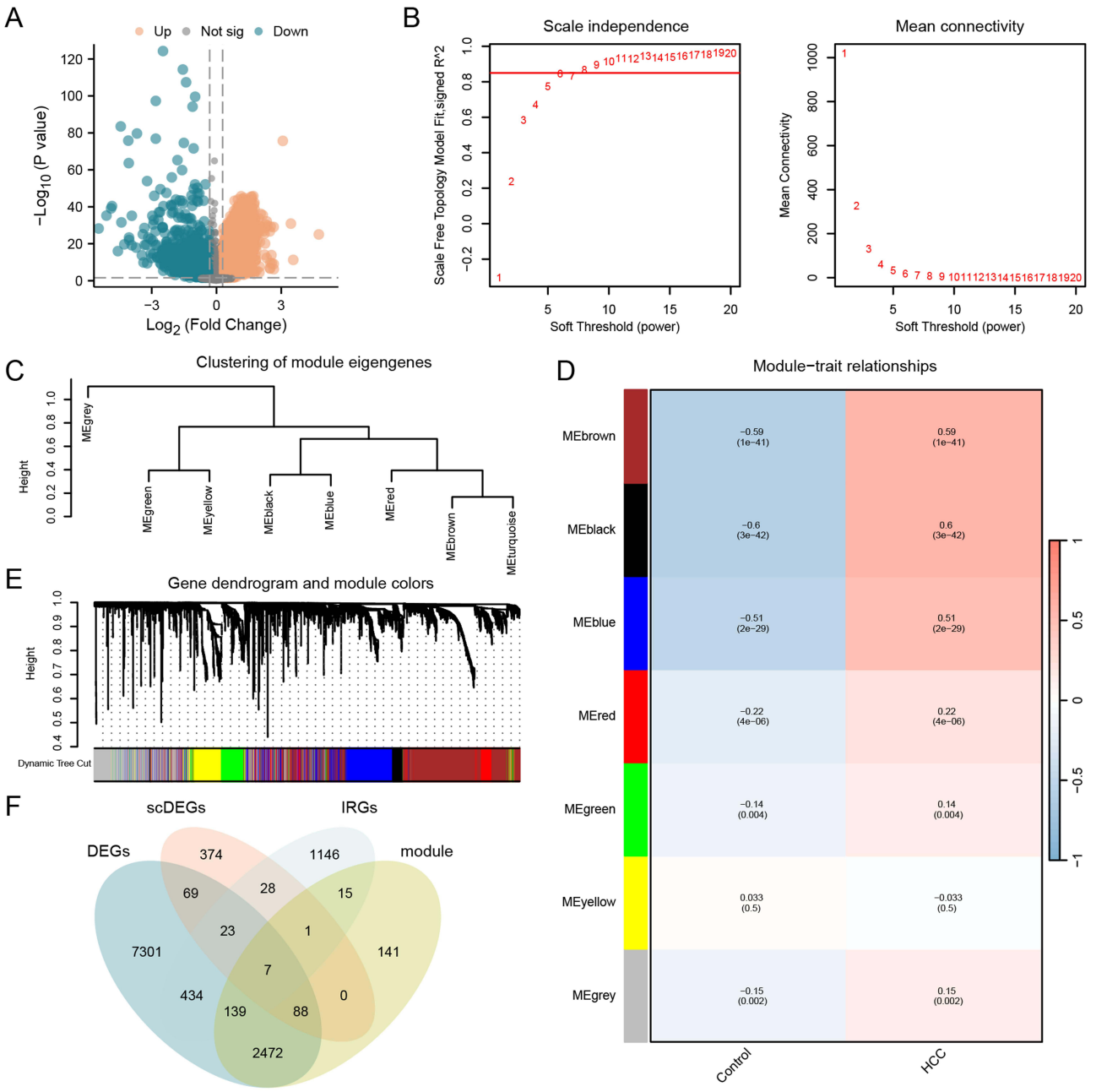

Figure 2 Identification of the hub genes (**A**) Volcano plot of differential gene expression analysis in the TCGA-LIHC. (**B**) Selection of the optimal soft-thresholding power and assessment of network connectivity. (**C**) Module clustering results for the top 5,000 genes based on MAD. (**D**) Correlation between module eigengenes and HCC/control groups. (**E**) Gene dendrogram and module clustering results. (**F**) Venn diagram of IRGs, scDEGs, DEGs, and module genes.

significantly associated with poor prognosis (Figure 5C–F). Moreover, time-dependent ROC analysis demonstrated that these markers had strong prognostic value at 1-, 3-, and 5-year intervals (Figure 5G–J).

Establishment of a Nomogram Integrating with Clinical Characteristics

Clinical characteristics from the TCGA-LIHC dataset were integrated with the predicted values of the SVM model for multivariate Cox regression analysis. Based on this analysis, a nomogram was constructed to estimate the survival probability of patients with HCC (Figure 6A). Among all the included variables, the predicted value of SVM demonstrated the strongest prognostic value, while the effect of gender was relatively weak. The calibration curve showed that

https://doi.org/10.2147/JHC.S574690

Table 4 The Parameters of the Five Machine Learning Algorithms for Feature Selection

Methods	Key Function	Parms
LASSO	Glmnet::cv.glmnet()	Family: "binomial" s(λ): "lambda.min" nfolds: 5 iter.times: 1000
LVQ	Caret::train()	trControl: trainControl(method="repeatedcv", number=5)
	Method = "lvq"	Preprocess: "scale" tuneGrid: data.frame(size = 6, k = 1:2) cutoff_line: 0.5
Bagged Trees (Treebag)	Caret::rfe()	Rfecontrol: rfeControl(functions=treebagFuncs, method="cv", number=3)
	Functions = treebagFuncs	Sizes: c(2:ncol(df))
Boruta	Boruta::Boruta()	Dotrace: 3
		Tentative: TentativeRoughFix()
Random Forest (RF)	Caret::rfe()	Rfecontrol: rfeControl(functions=rfFuncs, method="cv")
	Functions = rfFuncs	Sizes: c(2:ncol(df))

Abbreviations: LASSO, Least Absolute Shrinkage and Selection Operator; LVQ, Learning Vector Quantization; RFE, Recursive Feature Elimination; RF, Random Forest.

the nomogram exhibited optimal clinical predictive performance at the 5-year (Figure 6B). Additionally, DCA demonstrated that the net clinical benefit of the prognostic model followed the pattern: 5-year > 3-year > 1-year (Figure 6C–E).

Underlying Molecular Mechanisms of IDMS

To investigate the biological processes underlying the prognostic relevance of IDMS, we performed GSEA and GSVA using transcriptomic data. The results revealed that oncogenic signaling pathways—including PI3K_AKT_MTOR, NOTCH, and WNT_BETA—were significantly activated in the high-risk group (Figure 7A–E and Table S1). In contrast, the low-risk group exhibited marked alterations in bile acid and fatty acid metabolism (Figure 7F and G and Table S2). These findings indicate a strong association between HCC risk stratification and dysregulated biological processes and metabolic pathways relevant to cancer progression.

Analysis of Mutation and Immune Characteristics in Different Subgroups

SMs in IDMS were analyzed and visualized using "maftools" (Figure 8A). The results showed two types of SMs, with missense mutations being the most prevalent. In addition, single nucleotide polymorphisms represented the predominant mutation type, and C >T transitions were the most frequently observed single nucleotide variants in HCC samples. CNVs were present in all IDMS-related genes in the TCGA-LIHC samples (Figure 8B and C). Among them, S100A10 showed the highest frequency of CNV amplification and the lowest frequency of CNV deletion, particularly in the variant group compared to the non-variant group. We further evaluated immunotherapy sensitivity in HCC samples by comparing risk groups classified by SVM model. No statistically significant were observed in TIDE scores or TMB among risk subgroups. In contrast, the MSI score was significantly higher in the high-risk group (Figure 8D and F). MSI is an established biomarker for predicting tumor response to immunological therapy.[31] These findings indicate that patients in the high-risk group might have a greater likelihood of benefiting from immunotherapeutic strategies.

The Correlation Between the IDMS and Single-Cell Characteristics

A large amount of evidence shows an inseparable connection between the immune microen- vironment and HCC. In this study, the CIBERSORT algorithm was employed to quantify the abundance of 22 immune cell types and to compare their

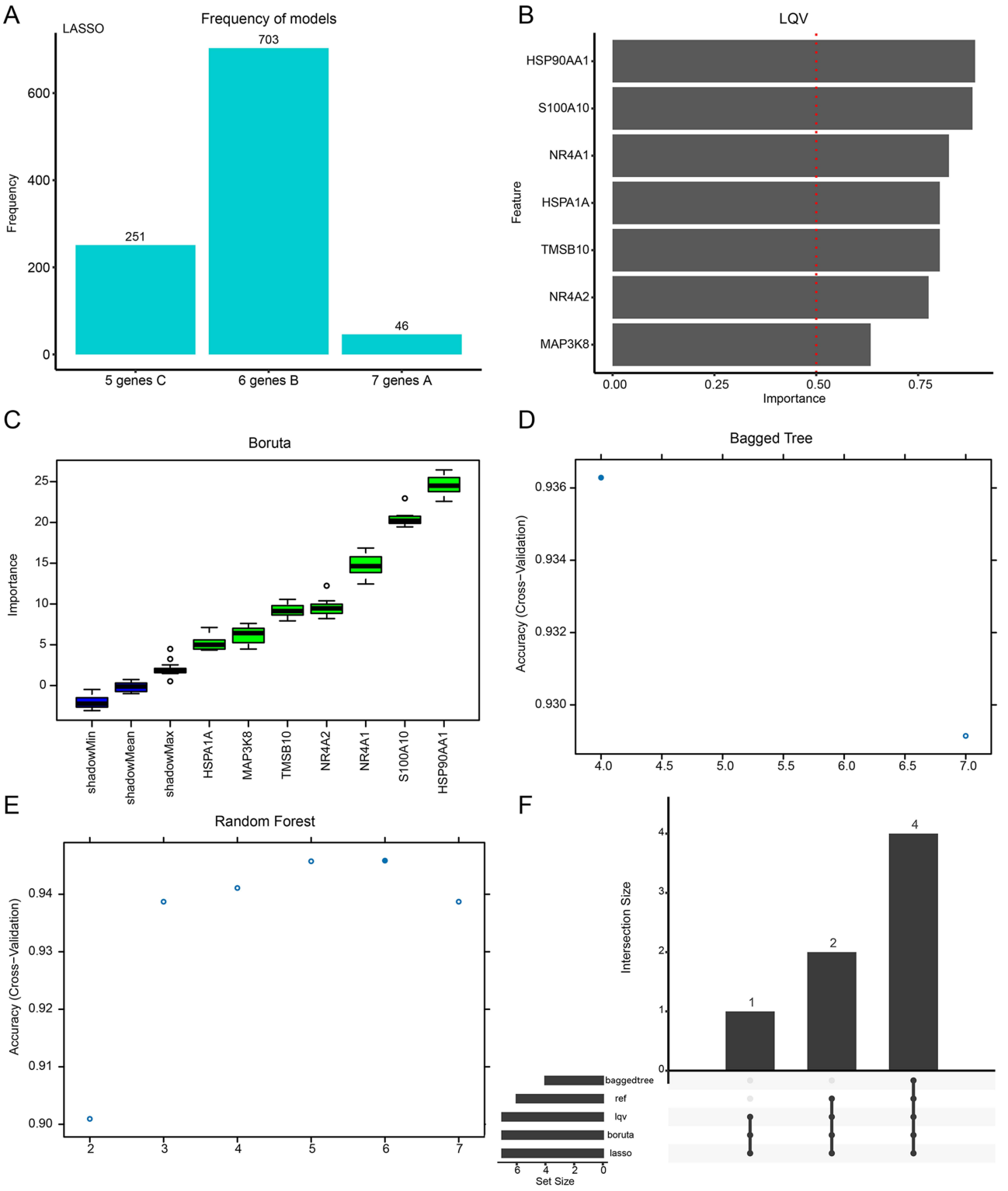

Figure 3 Screening immune-derived genes (**A**) Gene frequency distribution derived from the LASSO algorithm. (**B**) Genes selected using LVQ algorithm.The red dotted line indicating the threshold of importance for gene screening. (**C**) Gene selection using the Boruta algorithm. (**D**) Variable importance and model performance from the Bagged Trees. (**E**) Variable contribution and model accuracy from the random forest model. (**F**) The Upset diagram shows the intersection of the screening results of five machine learning algorithms and determines the final key genes.

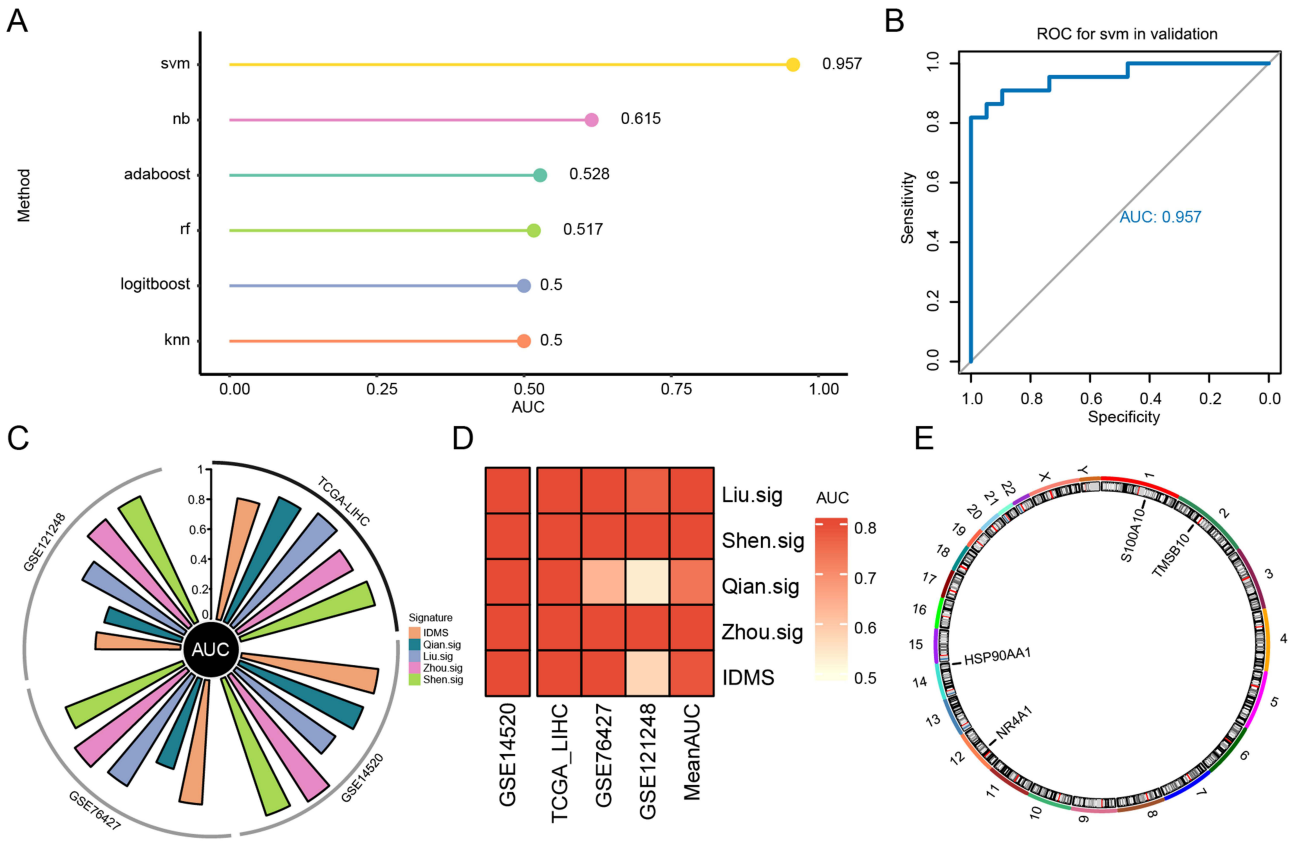

Figure 4 Identification of risk model and comparison with other HCC signatures (**A**) AUC values of risk models constructed using six machine learning algorithms. (**B**) ROC curve of the SVM model. (**C**) Circos plot comparing the performance of various published HCC signatures. (**D**) Heatmap showing predictive performance of the risk model across datasets. (**E**) Chromosomal mapping of IDMS.

distribution across risk subgroups (Figure 9A). Additionally, a more in-depth analysis of immune cell infiltration uncovered intricate correlations among various immune cell populations. For instance, Plasma cells are positively correlated with naïve B cells, whereas monocytes are negatively correlated with follicular helper T cells (Figure 9B). Interestingly, Tregs exhibited a significant negative correlation with NR4A1, while displaying significant positive correlations with S100A10 and TMSB10. Moreover, HSP90AA1 demonstrated a significant negative association with M1-type macrophages, suggesting a potential role of IDMS in immune tolerance mechanisms (Figure 9C).

Analysis of Drug Sensitivity and Molecular Docking

To predict potential therapeutic agents for high-risk patients, we carried out a comprehensive drug sensitivity analysis using DGIdb and the DrugBank database to screen for drugs or small-molecule compounds interacting with IDMS (Figure 10A and Table 5). In addition, candidate therapeutic compounds and their target genes were evaluated using the Connectivity Map (CMap) (Figure 10B and Table 6). Based on compound susceptibility scores, the top two drugs were selected for molecular docking. The results indicated that HSP90AA1 exhibited moderate binding affinity with geldanamycin (Figure 10C). This interaction involved amino acid residues TYR493, GLU527, VAL530, GLN531, LEU533, LYS534, THR540, LEU541, VAL542, SER543, LYS546, THR603, TYR604, GLY605, TRP606, THR607, and MET610, which engaged in hydrogen bonding, weak hydrogen bonding, and hydrophobic interactions. Furthermore, HSP90AA1 also demonstrated moderate binding affinity with alvespimycin (Figure 10D). The interacting residues included TYR493, VAL530, GLN531, LYS534, THR540, LEU541, VAL542, SER543, LYS546, THR603, TYR604, and GLY605, forming similar hydrogen bonds and hydrophobic contacts. Finally, we evaluated the expression of IDMS in five cell lines and found HSP90AA1, NR4A1, and S100A10 significantly upregulated in tumor cell lines versus the liver cell line (Figure 10E–H).

https://doi.org/10.2147/JHC.S574690

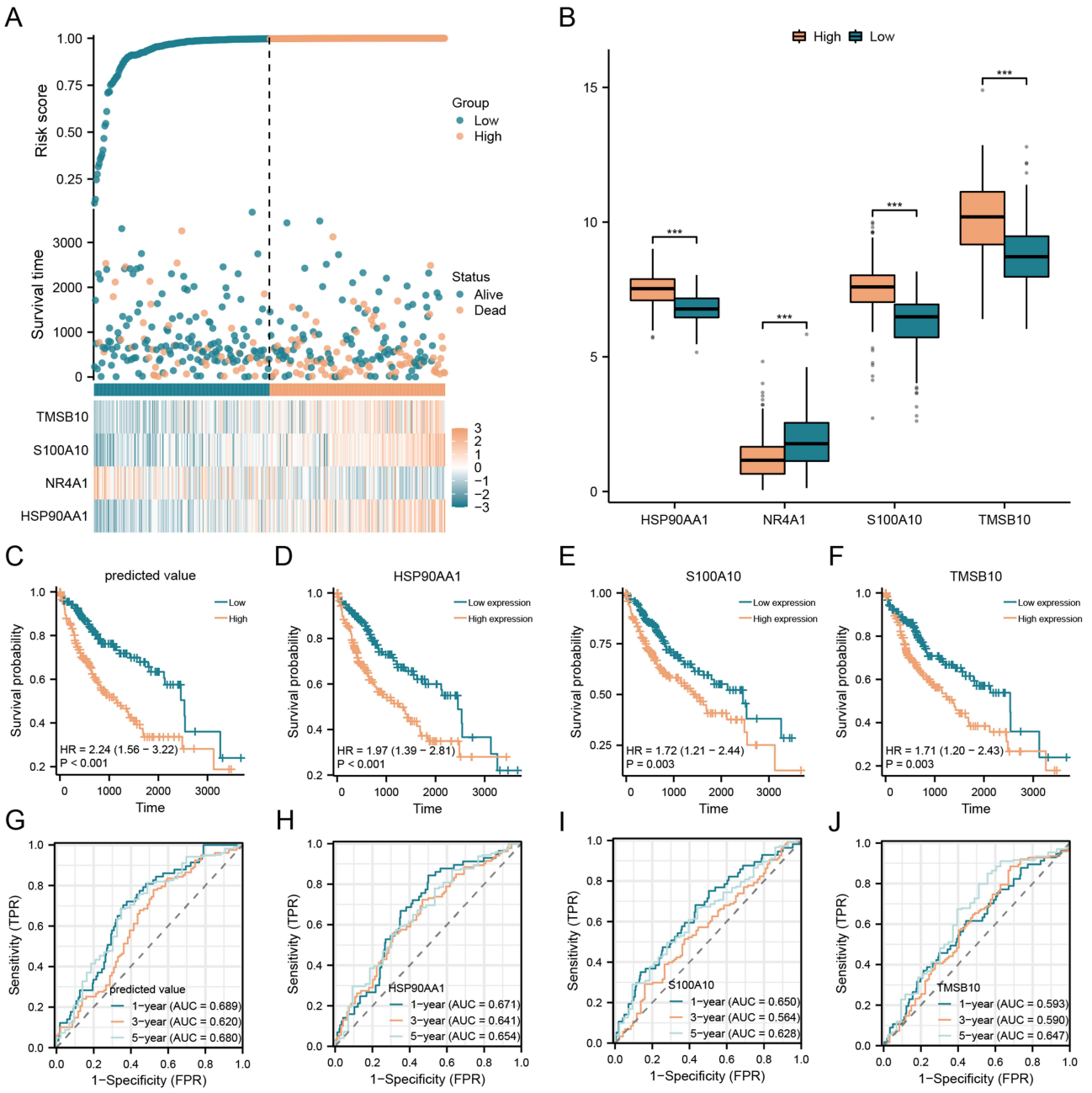

Figure 5 Prognostic performance of the SVM for HCC (**A**) Risk factor plot within risk subgroups. (**B**) Compare gene expression levels among risk subgroups. (**C-F**) Kaplan–Meier survival curves for predicted value (**C**), HSP90AA1 (**D**), S100A10 (**E**), and TMSB10 (**F**). (**G–J**) Time-dependent ROC curves for predicted value (**G**), HSP90AA1 (**H**), S100A10 (**I**), and TMSB10 (**J**). ***$p < 0.001$.

Discussion

HCC remains a major malignancy that significantly threatens human health. Previous studies investigating the molecular basis of HCC have identified numerous potential biomarkers for diagnosis and prevention.[32–34] For instance, alpha-fetoprotein (AFP) is a commonly used biomarker for HCC diagnosis, however, its limited sensitivity and specificity restrict its clinical utility.[35] In addition, vascular endothelial growth factor A has demonstrated potential for predicting the efficacy of sorafenib treatment.[36] Promising molecular markers have also been identified in circulating tumor DNA (ctDNA) and circulating tumor cells.[37] Nevertheless, due to a lack of large-scale prospective studies, robust molecular biomarkers for clinical decision-making in HCC remain limited. At present, pathological classification remains a primary tool for determining treatment strategies. However, substantial differences in patient prognosis are frequently observed

https://doi.org/10.2147/JHC.S574690

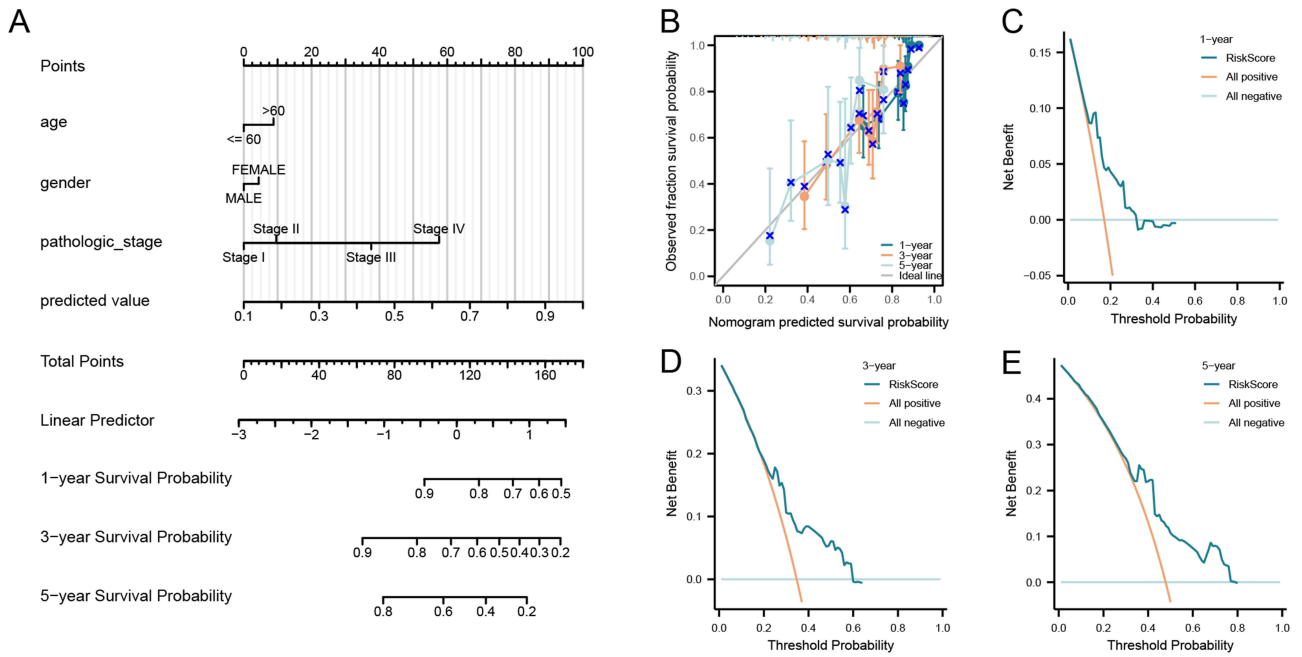

Figure 6 Establishment and evaluation of the nomogram (**A**) Nomogram integrating predicted value with clinical characteristics. (**B**) Calibration curves evaluating 1-, 3-, and 5-year survival predictions. (**C–E**) DCA of the prognostic model at 1-year (**C**), 3-year (**D**), and 5-year (**E**).

under this model. Therefore, in the context of HCC—characterized by limited diagnostic and therapeutic options and marked heterogeneity—identifying reliable molecular biomarkers to support the implementation of PPPM is a critical and urgent need.

Advancements in multi-omics technologies have facilitated the elucidation of tumor molecular mechanisms.[38] In this work, we integrated bulk and single-cell transcriptome datasets to explore potential biomarkers and characteristics of immune cell infiltration in HCC. To minimize bias stemming from a single algorithm, this study applied five machine learning algorithms with cross-validation to identify four immune-derived genes that exhibit greater stability and biological significance. Subsequently, we comprehensively analyzed the predictive performance of six machine learning approaches and selected the optimal risk model. Compared with the previous immune-related prognostic models constructed based on single-cell transcriptomes and bulk RNA-seq,[39,40] through cross-validation with multiple algorithms and independent validation across multiple cohorts, immune-derived molecular features with higher generalization ability were obtained.More importantly, survival analysis and nomogram indicated that IDMS exhibits a robust capacity for risk stratification and prognostic value. This result indicates that IDMS not only has better predictive performance in the prognosis of HCC, but can also more effectively identify high-risk patients and guiding the development of targeted surveillance and intervention strategies. For high-risk individuals, clinicians may consider increasing monitoring frequency and recommending lifestyle interventions such as routine liver ultrasounds, smoking cessation, and reduced alcohol intake. Furthermore, detecting ctDNA can enhance the early diagnosis of HCC and prevents its occurrence.[41]

In the past few years, immunotherapy has become a major focus of basic and clinical oncology research, with promising advances in its application to HCC.[42,43] However, challenges persist, including low objective response rates and adverse treatment effects.[44] Therefore, identifying patients most likely to benefit from immunotherapy is a key strategy for improving outcomes and personalizing treatment plans. Multi-omics and drug sensitivity analysis in this study indicated that patients in the high-risk group were more likely to benefit, potentially reducing overall treatment costs. Moreover, geldanamycin and alvespimycin may exhibit superior therapeutic efficacy. Geldanamycin inhibits the function of HSP90AA1, leading to the degradation of multiple tumor-related proteins, thereby preventing the proliferation of tumor cells and inducing apoptosis.[45,46] Alvespimycin inhibits tumor proliferation, metastasis and angiogenesis by suppressing pathways such as PI3K/AKT and MAPK, and enhances the sensitivity to chemotherapy.[47–49]Taken together,

https://doi.org/10.2147/JHC.S574690

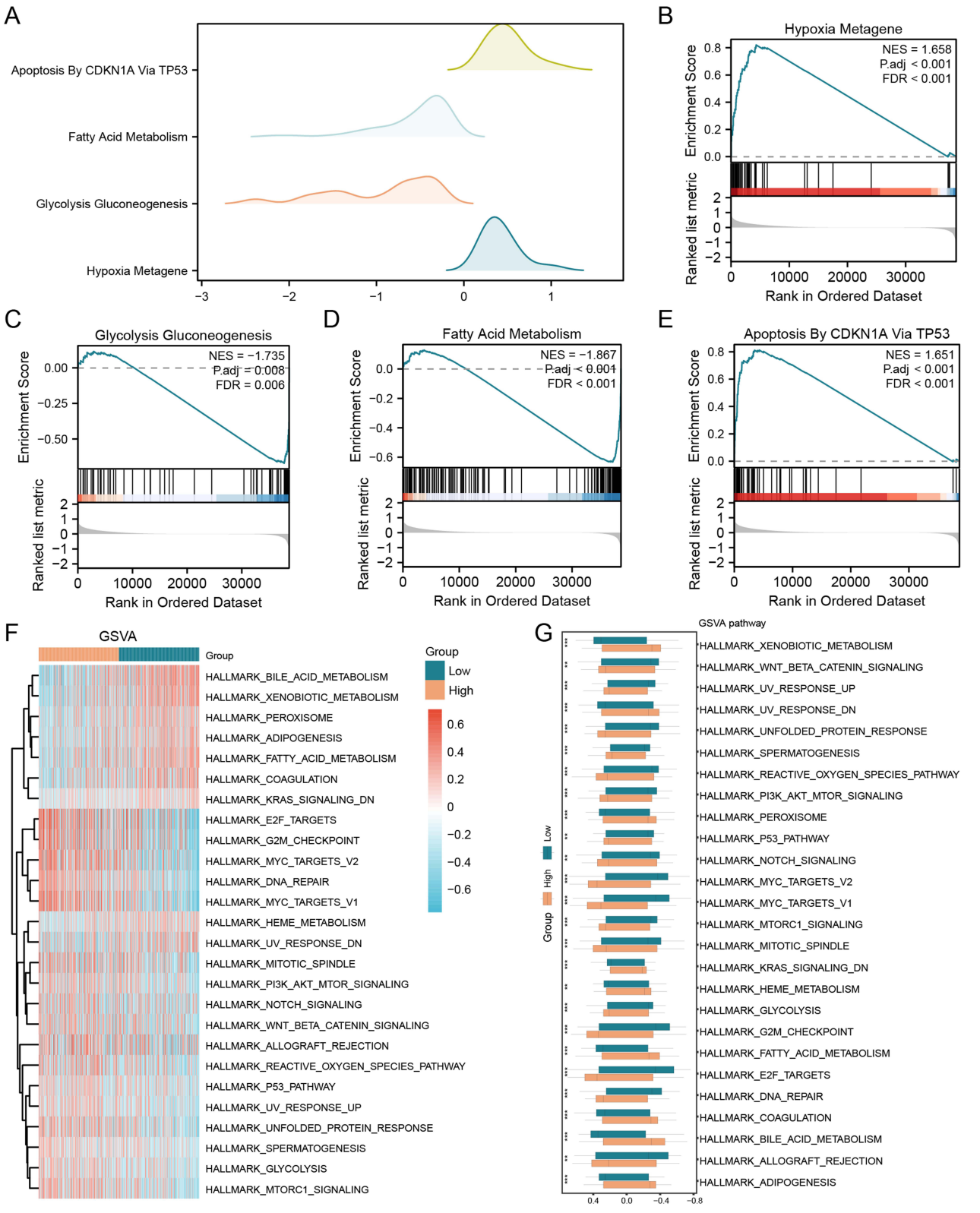

Figure 7 Transcriptome characteristics of the subgroups (**A**) Ridge plot of biological functions identified by GSEA. (**B**–**E**) GSEA enrichment plots showing upregulation of hypoxia metagene (**B**), glycolysis/gluconeogenesis (**C**), fatty acid metabolism (**D**), and CDKN1A-mediated apoptosis via TP53 (**E**).(**F** and **G**) Heatmap (**F**) and box plot (**G**) showing GSVA results comparing risk subgroups.

https://doi.org/10.2147/JHC.S574690

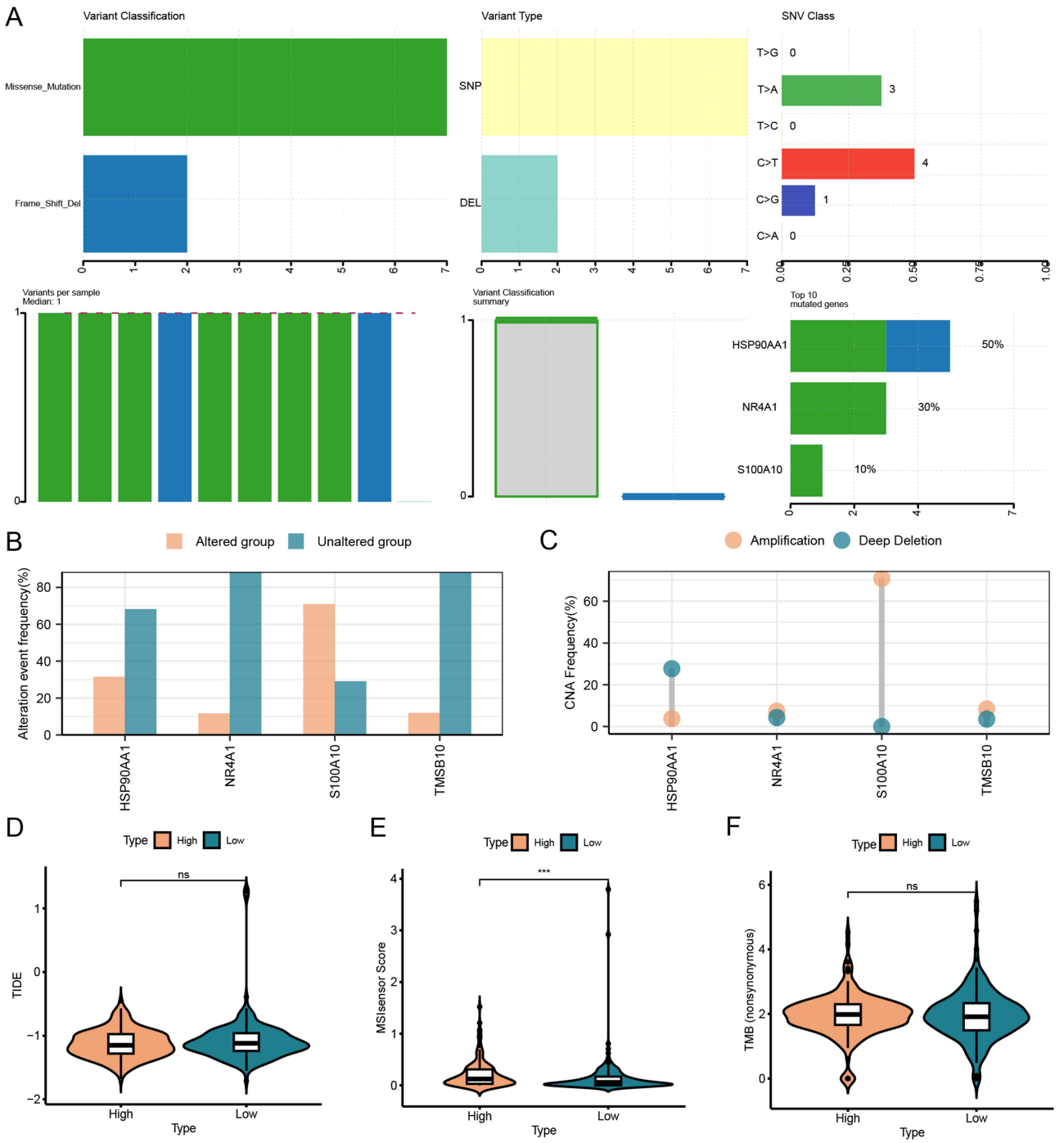

Figure 8 Mutation and immune characteristics in IDMS and its subgroups. (**A**) SM landscape of IDMS in the TCGA-LIHC. (**B** and **C**) CNV profiles of IDMS in TCGA-LIHC. (**D**–**F**) Comparison of immunotherapy response indicators between high- and low-risk groups: (**D**) TIDE scores, (**E**) MSI scores, and (**F**) TMB scores.

these findings suggest that IDMS are superior to previous markers in the prediction and personalized treatment of HCC, and are more conducive to the clinical management of PPPM for HCC patients.

Through the analysis of heterogeneity of IDMS at the Inter-cellular level, we found that all genes included in the IDMS are implicated in immune evasion or tumorigenesis. Tregs are a subset of T cells with immunosuppressive functions. Within the tumor microenvironment (TME) of HCC, persistent infiltration of Tregs promotes an immunosuppressive niche, thereby facilitating tumor progression and immune escape.[18,50,51] NR4A1, a transcription factor, has been

https://doi.org/10.2147/JHC.S574690

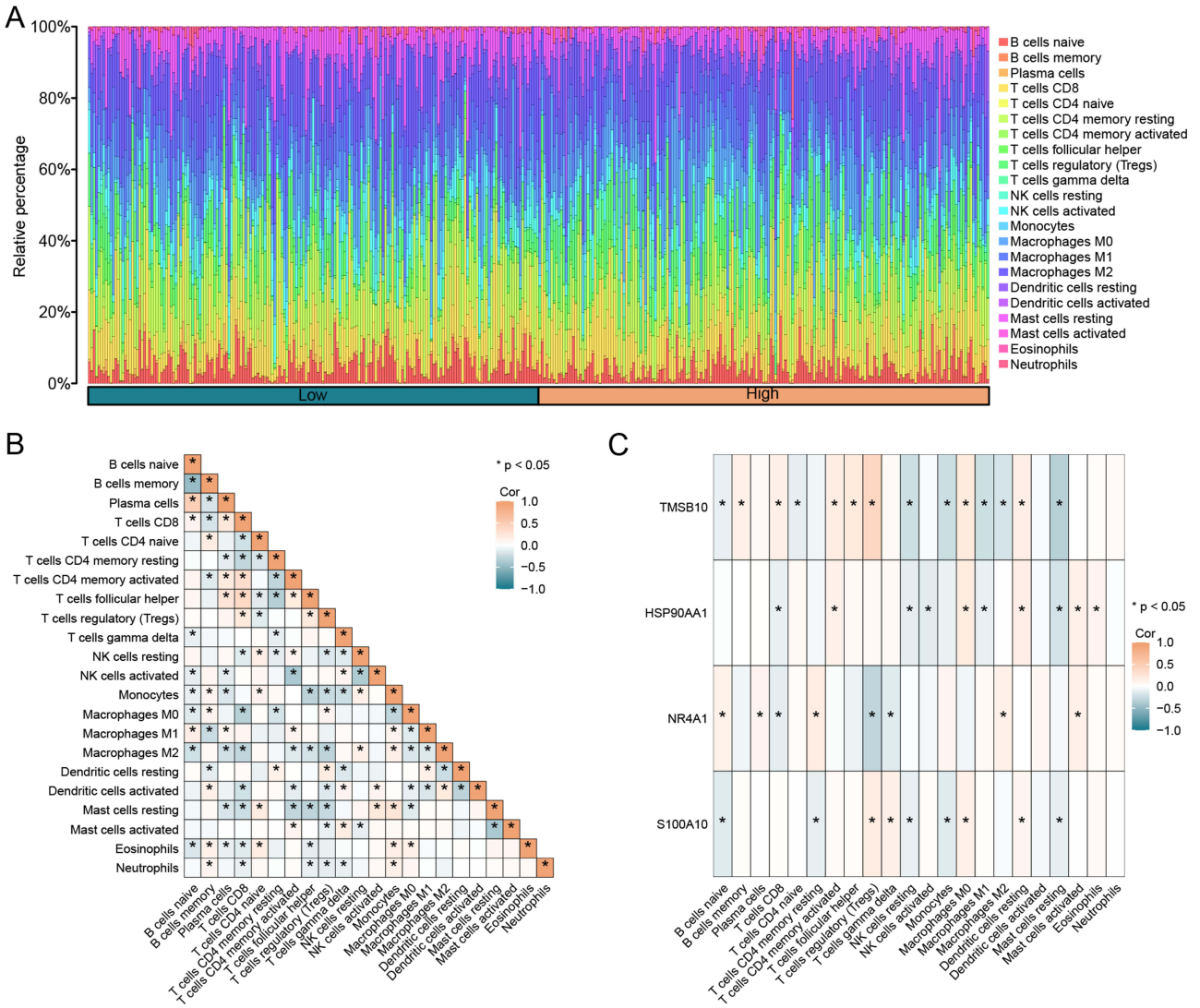

Figure 9 The correlation between the IDMS and single-cell characteristics (**A**) Stacked bar chart showing the distribution of immune cell types across risk subgroups. (**B**) Heatmap illustrating correlations among immune cell infiltration levels. (**C**) Heatmap showing correlations between immune cell abundance and IDMS expression levels.*p <0.05.

shown to promote the differentiation of naïve T cells into Tregs and enhance immunosuppression by modulating signaling pathways downstream of the T cell receptor (TCR).[52,53] S100A10 has been reported by Wang et al to activate the cPLA2 and 5-LOX axis, leading to $CD8^+$ T cell exhaustion in HCC and promoting sustained tumor progression and metastasis.[54] TMSB10, a member of the β-thymosin family, is overexpressed in various cancers.[55,56] Knockdown of TMSB10 has been shown to facilitate the polarization of macrophages towards the M1 phenotype, thereby improving antitumor immunity.[57] HSP90AA1, a molecular chaperone, interacts with numerous client proteins and is involved in multiple biological processes related to cancer. Notably, plasma levels of HSP90AA1 are significantly elevated in HCC patients and exhibit superior diagnostic performance compared to AFP, with strong associations with tumor progression and metastasis.[58,59] Importantly, both S100A10 and TMSB10 are positively correlated with Treg abundance and are highly expressed in HCC tissues, suggesting that they may play key roles in mediating Treg-associated immune escape mechanisms. However, the molecular mechanisms involved are not yet fully elucidated. Moreover, the potential biases that may arise from the publicly available data also require further investigation.

We investigated the underlying mechanisms of the IDMS using multi-omics analysis. DEGs between the risk subgroups were primarily correlated with cell signaling pathways—such as the PI3K, WNT, and MYC pathways—as

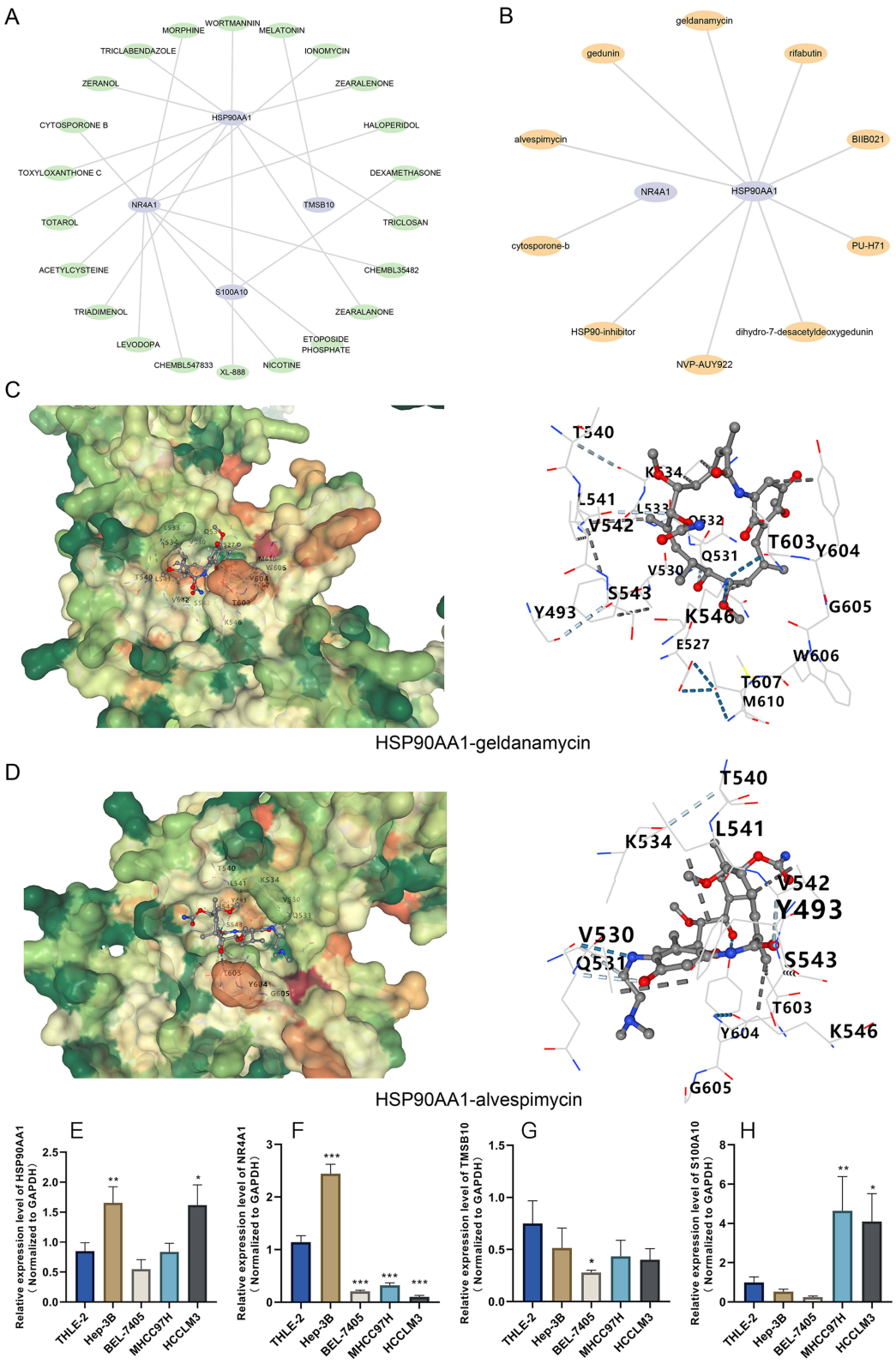

Figure 10 Drug sensitivity and molecular docking analysis of IDMS (**A**) Interaction network between IDMS and candidate drugs or small-molecule compounds based on DGIdb and DrugBank analysis. (**B**) Predicted interactions between NR4A1, HSP90AA1, and candidate drugs based on CMap analysis. (**C–D**) Molecular docking results of HSP90AA1 with geldanamycin and alvespimycin, showing overall docking structure and interaction strength diagrams from left to right. The protein is color-coded from green (hydrophilic) to red (hydrophobic) to reflect amino acid properties. Hydrogen bonds are indicated by dashed blue lines; weak hydrogen bonds by light blue dashed lines; and hydrophobic interactions by gray dashed lines. (**E–H**) Validation of mRNA expression levels of HSP90AA1 (**E**), NR4A1 (**F**), TMSB10 (**G**), and S100A10 (**H**) in one liver cell line (THLE-2) and four HCC cell lines (Hep-3B, BEL-7405, MHCC97H, and HCCLM3) via qRT-PCR.

https://doi.org/10.2147/JHC.S574690

Table 5 List of Drugs Related IDMS From DGIdb Database

Gene	Drug
HSP90AA1	TRICLABENDAZOLE
HSP90AA1	TOXYLOXANTHONE C
HSP90AA1	TRIADIMENOL
HSP90AA1	XL-888
HSP90AA1	ZEARALANONE
HSP90AA1	TRICLOSAN
HSP90AA1	ZEARALENONE
HSP90AA1	WORTMANNIN
HSP90AA1	ZERANOL
HSP90AA1	TOTAROL
NR4A1	LEVODOPA
NR4A1	NICOTINE
NR4A1	CHEMBL35482
NR4A1	HALOPERIDOL
NR4A1	IONOMYCIN
NR4A1	MORPHINE
NR4A1	CYTOSPORONE B
NR4A1	ACETYLCYSTEINE
NR4A1	CHEMBL547833
NR4A1	ETOPOSIDE PHOSPHATE
S100A10	DEXAMETHASONE
TMSB10	MELATONIN

Table 6 List of Drug-Related IDMS and Drug Susceptibility Scores in CMap Database

Target	Name	Score
HSP90AA1	Geldanamycin	90.43
HSP90AA1	Alvespimycin	84.94
HSP90AA1	HSP90-inhibitor	83.02
HSP90AA1	NVP-AUY922	58.76
HSP90AA1	Dihydro-7-desacetyldeoxygedunin	52.11
HSP90AA1	PU-H71	30.86
HSP90AA1	BIIB021	27.64
HSP90AA1	Rifabutin	9.57
HSP90AA1	Gedunin	−22.02
NR4A1	Cytosporone-b	−43.98

well as metabolic processes including glycolysis, xenobiotic metabolism, and fatty acid metabolism. These results may partially explain the differences in prognosis observed between IDMS-defined subgroups. Furthermore, metabolism-associated biological pathways may represent promising directions for future research, and the development of targeted inhibitors against these pathways could offer therapeutic benefit in preventing HCC progression.

Nevertheless, there are certain limitations in this investigation. First of all, all data were sourced from publicly databases. Although the immune-derived molecular signature evaluated in this study demonstrated consistent prognostic predictive performance across multiple independent datasets, their clinical utility requires further validation in large-scale, prospective cohort studies. Additionally, our findings are derived from computational bioinformatics analyses, and functional experimental validation is lacking to confirm the biological roles of the genes comprising the IDMS. We hope

that the publication of this study will attract broader research interest in the IDMS and contribute to a deeper understanding of its function in the pathogenesis and progression of HCC.

Conclusion

We identified an IDMS that shows potential as a valuable tool for predictive, preventive, and personalized medicine in HCC. Through multi-omics analysis, this study may offer novel perspectives on the molecular mechanisms that underlie the occurrence and development of HCC.

Data Sharing Statement

The raw data in this study are available from the following databases: TCGA database (https://portal.gdc.cancer.gov/), GEO database (http://www.ncbi.nlm.nih.gov/geo), ImmPort database (https://www.immport.org/), DGIdb (https://dgidb.org/), DrugBank database (https://go.drugbank.com/). The parameters used for machine learning model are listed in Table 4. If needed, please refer to the corresponding author for further details.

Ethical Approval

According to Article 32 of the *Ethical Review Measures for Human Life Science and Medical Research* in China (issued on February 18, 2023), the data for this study were obtained from the public databases and fall within the scope of ethical review exemption.

Consent to Participate

All authors volunteered to participate in this study.

Consent for Publication

All authors have provided their consent for the publication of this paper.

Funding

Funding for this research was provided by the Joint Funds for the Innovation of Science and Technology, Fujian province (No.2021Y9064) and Fujian Provincial Natural Science Foundation of China (No.2022J01730).

Disclosure

The authors report no conflicts of interest in this work.

References

1. Villanueva A. *Hepatocellular Carcinoma N Engl J Med*. 2019;380(15):1450–1462.
2. McGiynn KA, Petrick JL, EL-Serag HB. Epidemiology of hepatocellular carcinoma. Hepatology. 2021;73:4–13.
3. Allemani C, Matsuda T, Di Carlo V, et al. Global surveillance of trends in cancer survival 2000-14 (Concord-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet*. 2018;391(10125):1023–1075. doi:10.1016/S0140-6736(17)33326-3
4. Golabi P, Fazel S, Otgonsuren M, et al. Mortality assessment of patients with Hepatocellularcarcinoma according to underlying disease and treatment modalities. *Medicine*. 2017;96(9):e5904. doi:10.1097/MD.0000000000005904
5. Lu M, Zhan H, Liu B, et al. N6-methyladenosine-related non-coding RNAs are potential prognostic and immunotherapeutic responsiveness biomarkers for bladder cancer. *EPMAJ*. 2021;12(4):589–604. doi:10.1007/s13167-021-00259-w
6. He S, Qiao J, Wang L, et al. A novel immune-related gene signature predicts the prognosis of hepatocellular carcinoma. *Front Oncol*. 2022;12:955192. doi:10.3389/fonc.2022.955192
7. Xin C, Huang B, Chen M, et al. Construction and validation of an immune-related, LncRNA prognostic model for hepatocellular carcinoma. *Cytokine*. 2022;156:155923. doi:10.1016/j.cyto.2022.155923
8. Zhang G, Su L, Lv X, et al. A novel tumor doubling time-related immune gene signature for prognosis prediction in hepatocellular carcinoma. *Cancer Cell Int*. 2021;21(1):522. doi:10.1186/s12935-021-02227-w
9. Beumer BR, Buettner S, Galjart B, et al. Systematic review and meta-analysis of validated prognostic models for resected hepatocellular carcinoma patients. *Eur J Surg Oncol*. 2022;48(3):492–499. doi:10.1016/j.ejso.2021.09.012
10. Han Y, Song L, Lv L, et al. Unraveling the heterogeneity of tumor immune microenvironment in hepatocellular carcinoma by singlecell rna sequencing and its implications for prognosis and therapeutic response. *Turk J Gastroenterol*. 2024;35(12):876–888. doi:10.5152/tjg.2024.24118

11. Zucman-Rossi J, Villanueva A, Nault JC, et al. Genetic landscape and biomarkers ofhepatocellular carcinoma. *Gastroenterology*. 2015;149 (5):1226–1239.e4. doi:10.1053/j.gastro.2015.05.061
12. Nakagawa H, Fujita M, Fujimoto A. Genome sequencing analysis of liver cancer for precision medicine. *Semin Cancer Biol*. 2019;55:120–127. doi:10.1016/j.semcancer.2018.03.004
13. Bingzhe L, Yunpeng W, Dongjiang M, et al. Immunotherapy: reshape the tumor immunemicroenvironment. *Front Immunol*. 2022;13:844142. doi:10.3389/fimmu.2022.844142
14. Qin X, Chen Y, Ma S, et al. Immune-related gene TM4SF18 could promote the metastasis of gastric cancer cells and predict the prognosis of gastric cancer patients. *Mol Oncol*. 2022;16(22):4043–4059. doi:10.1002/1878-0261.13321
15. Zhuang Y, Li S, Liu C, et al. Identification of an individualized immune-related prognostic risk score in lung squamous cell cancer. *Front Oncol*. 2021;11:546455. doi:10.3389/fonc.2021.546455
16. Chen H, Luo J, Guo J. Construction and validation of a 7-immune gene model for prognostic assessment of esophageal carcinoma. *Med Sci Monit*. 2020;26:e927392. doi:10.12659/MSM.927392
17. Zheng L, Qin S, Si W, et al. Pan-cancer single-cell landscape of tumor-infiltrating T cells. *Science*. 2021;374(6574):abe6474. doi:10.1126/science.abe6474
18. Zheng C, Zheng L, Yoo JK, et al. Landscape of infiltrating t cells in liver cancer revealed by single-cell sequencing. *Cell*. 2017;169(7):1342–1356. doi:10.1016/j.cell.2017.05.035
19. Obeid M, Tesniere A, Ghiringhelli F, et al. Calreticulin exposure dictates the immunogenicity of cancer cell death. *Nat Med*. 2007;13(1):54–61. doi:10.1038/nm1523
20. Yu WD, Sun G, Li J, et al. Mechanisms and therapeutic potentials of cancer immunotherapy in combination with radiotherapy and/or chemotherapy. *Cancer Lett*. 2019;28(452):66–70. doi:10.1016/j.canlet.2019.02.048
21. Salas-Benito D, Pérez-Gracia JL, Ponz-Sarvisé M, et al. Paradigms on immunotherapy combinations with chemotherapy. *Cancer Discov*. 2021;11 (6):1353–1367. doi:10.1158/2159-8290.CD-20-1312
22. Feng Q, Huang Z, Song L, et al. Combining bulk and single-cell RNA-sequencing data todevelop an NK cell-related prognostic signature for hepatocellular carcinoma based onan integrated machine learning framework. *Eur J Med Res*. 2023;28(1):306. doi:10.1186/s40001-023-01300-6
23. Roessler S, Jia HL, Budhu A, et al. A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. *Cancer Res*. 2010;70(24):10202–10212. doi:10.1158/0008-5472.CAN-10-2607
24. Grinchuk OV, Yenamandra SP, Iyer R, et al. Tumor-adjacent tissue co-expression profile analysis reveals pro-oncogenic ribosomal gene signature for prognosis of resectable hepatocellular carcinoma. *Mol Oncol*. 2018;12(1):89–113. doi:10.1002/1878-0261.12153
25. Wang SM, Ooi LL, Hui KM. Identification and validation of a novel gene signature associated with the recurrence of human hepatocellular carcinoma. *Clin Cancer Res*. 2007;13(21):6275–6283. doi:10.1158/1078-0432.CCR-06-2236
26. Langfelder P, Horvath S. WGCNA:an R package for weighted correlation network analysis. *BMC Bioinf*. 2008;29(9):559. doi:10.1186/1471-2105-9-559
27. Mermel CH, Schumacher SE, Hill B, et al. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol*. 2011;12(4):R41. doi:10.1186/gb-2011-12-4-r41
28. Liu GM, Zeng HD, Zhang CY, et al. Identification of a six-gene signature predicting overall survival for hepatocellular carcinoma. *Cancer Cell Int*. 2019;19:138. doi:10.1186/s12935-019-0858-2
29. Zhou Y, Li X, Long G, et al. Identification and validation of a tyrosine metabolism-related prognostic prediction model and characterization of the tumor microenvironmentinfiltration in hepatocellular carcinoma. *Front Immunol*. 2022;13:994259. doi:10.3389/fimmu.2022.994259
30. Shen B, Zhang G, Liu Y, et al. Identification and analysis of immune-related gene signature in hepatocellular carcinoma. *Genes*. 2022;13(10):1834. doi:10.3390/genes13101834
31. Yamamoto H, Watanabe Y, Arai H, et al. Microsatellite instability: a 2024 update. *Cancer Sci*. 2024;115(6):1738–1748. doi:10.1111/cas.16160
32. Wang D, Zhang L, Sun Z, et al. A radiomics signature associated with underlying gene expression pattern for the prediction of prognosis and treatment response in hepatocellular carcinoma. *Eur J Radiol*. 2023;167:111086. doi:10.1016/j.ejrad.2023.111086
33. Wang Y, Deng B. Hepatocellular carcinoma: molecular mechanism, targeted therapy, and biomarkers. *Cancer Metastasis Rev*. 2023;42(3):629–652. doi:10.1007/s10555-023-10084-4
34. Zhang ZM, Tan JX, Wang F, et al. Early diagnosis of hepatocellular carcinoma using machine learning method. *Front Bioeng Biotechnol*. 2020;8:254. doi:10.3389/fbioe.2020.00254
35. Ma S, Chan KW, Hu L, et al. Identification and characterizationof tumorigenic liver cancer stem/progenitor cells. *Gastroenterology*. 2007;132 (7):2542–2556. doi:10.1053/j.gastro.2007.04.025
36. Horwitz E, Stein I, Andreozzi M, et al. Human and mouse VEGFA-amplified hepatocellular carcinomas are highly sensitive to sorafenib treatment. *Cancer Discov*. 2014;4(6):730–743. doi:10.1158/2159-8290.CD-13-0782
37. Johnson P, Zhou Q, Dao DY, et al. Circulating biomarkers in the diagnosis and management of hepatocellular carcinoma. *Nat Rev Gastroenterol Hepatol*. 2022;19(10):670–681. doi:10.1038/s41575-022-00620-y
38. Lu M, Zhan X. The crucial role of multiomic approach in cancer research and clinically relevant outcomes. *EPMAJ*. 2018;9(1):77–102.
39. Li B, Zeng T, Chen C, et al. Unraveling the potential mechanism and prognostic value of pentose phosphate pathway in hepatocellular carcinoma: a comprehensive analysis integrating bulk transcriptomics and single-cell sequencing data. *Funct Integr Genomics*. 2025;25(1):11. doi:10.1007/s10142-024-01521-w
40. Chen Q, Zhang C, Meng T, et al. Prediction of clinical prognosis and drug sensitivity in hepatocellular carcinoma through the combination of multiple cell death pathways. *Cell Biol Int*. 2024;48(12):1816–1835. doi:10.1002/cbin.12235
41. Ye Q, Ling S, Zheng S, et al. Liquid biopsy in hepatocellular carcinoma: circulating tumor cells and circulating tumor DNA. *Mol Cancer*. 2019;18 (1):114. doi:10.1186/s12943-019-1043-x
42. Shimizu Y, Suzuki T, Yoshikawa T, et al. Next-generation cancer immunotherapy targeting glypican-3. *Front Oncol*. 2019;9:248. doi:10.3389/fonc.2019.00248
43. Obeid JM, Kunk PR, Zaydfudim VM, et al. Immunotherapy for hepatocellular carcinoma patients: is it ready for prime time? *Cancer Immunol Immunother*. 2018;67(2):161–174. doi:10.1007/s00262-017-2082-z

44. Pinter M, Jain RK, Duda DG. The current landscape of immune checkpoint blockade in hepatocellular carcinoma: a review. *JAMA Oncol*. 2021;7(1):113–123. doi:10.1001/jamaoncol.2020.3381
45. Talaei S, Mellatyar H, Asadi A, et al. Spotlight on 17-AAG 17-AAGas an Hsp90 inhibitor for molecular targeted cancer treatment. *Chem Biol Drug Des*. 2019;93(5):760–786. doi:10.1111/cbdd.13486
46. Wang Z, Fan L, Xu H, et al. HSP90AA1 is an unfavorable prognostic factor for hepatocellular carcinoma and contributes to tumorigenesis and chemotherapy resistance. *Transl Oncol*. 2024;50:102148. doi:10.1016/j.tranon.2024.102148
47. Wang Y, Li S, Ren T, et al. Mechanism of emodin in treating hepatitis B virus-associated hepatocellular carcinoma: network pharmacology and cell experiments. *Front Cell Infect Microbiol*. 2024;14:1458913. doi:10.3389/fcimb.2024.1458913
48. Liu Z, Zhang H, Yao J. Metabolomic profiling and network toxicology: mechanistic. insights into effect of gossypol acetate isomers in uterine fibroids and liver injury. *Pharmaceuticals*. 2024;17(10):1363. doi:10.3390/ph17101363
49. Sadaqat M, Qasim M, Tahir Ul Qamar M, et al. Advanced network pharmacology study reveals multi-pathway and multi-gene regulatory molecular mechanism of Bacopa monnieri in liver cancer based on data mining, molecular modeling, and microarray data analysis. *Comput Biol Med*. 2023;161:107059. doi:10.1016/j.compbiomed.2023.107059
50. Wang H, Zhang H, Wang Y, et al. Regulatory T cell and neutrophil extracellular trap. interaction contributes to carcinogenesis in non-alcoholic steatohepatitis. *J Hepatol*. 2021;75(6):1271–1283. doi:10.1016/j.jhep.2021.07.032
51. Gao Y, You M, Fu J, et al. Intratumoral stem-like CCR4+ regulatory T cells orchestrate the immunosuppressive microenvironment in HCC associated with hepatitis B. *J Hepatol*. 2022;76(1):148–159. doi:10.1016/j.jhep.2021.08.029
52. You M, Gao Y, Fu J, et al. Epigenetic regulation of HBV-specific tumor-infiltrating T cells in HBV-related. *HCC.Hepatology*. 2023;78(3):943–958. doi:10.1097/HEP.0000000000000369
53. Hiwa R, Brooks JF, Mueller JL, et al. NR4A nuclear receptors in T and B lymphocytes: gatekeepers of immune tolerance. *Immunol Rev*. 2022;307(1):116–133. doi:10.1111/imr.13072
54. Wang G, Shen X, Jin W, et al. Elucidating the role of S100A10 in CD8+ T cell exhaustion and HCC immune escape via the cPLA2 and 5-LOX axis. *Cell Death Dis*. 2024;15(8):573. doi:10.1038/s41419-024-06895-0
55. Zhang X, Ren D, Guo L, et al. Thymosin beta 10 is a key regulator of tumorigenesisand metastasis and a novel serum marker in breast cancer. *Breast Cancer Res*. 2017;19(1):15. doi:10.1186/s13058-016-0785-2
56. Lee SM, Na YK, Hong HS, et al. Hypomethylation of the thymosin β(10) gene is not associated with its overexpression in non-small cell lung cancer. *Mol Cells*. 2011;32(4):343–348. doi:10.1007/s10059-011-0073-z
57. Zeng J, Yang X, Yang L, et al. Thymosin β10 promotes tumor-associated macrophages M2 conversion and proliferation via the PI3K/Akt pathway in lung adenocarcinoma. *Respir Res*. 2020;21(1):328. doi:10.1186/s12931-020-01587-7
58. Fu Y, Xu X, Huang D, et al. Plasma heat shock protein 90alpha as a biomarker for the diagnosis of liver cancer: an official, large-scale, and multicenter clinical trial. *EBioMedicine*. 2017;24:56–63. doi:10.1016/j.ebiom.2017.09.007
59. Shi W, Feng L, Dong S, et al. FBXL6 governs c-MYC to promote hepatocellular carcinoma through ubiquitination and stabilization of HSP90AA1. *Cell Commun Signal*. 2020;18(1):100. doi:10.1186/s12964-020-00604-y

Journal of Hepatocellular Carcinoma

Dovepress
Taylor & Francis Group

Publish your work in this journal

The Journal of Hepatocellular Carcinoma is an international, peer-reviewed, open access journal that offers a platform for the dissemination and study of clinical, translational and basic research findings in this rapidly developing field. Development in areas including, but not limited to, epidemiology, vaccination, hepatitis therapy, pathology and molecular tumor classification and prognostication are all considered for publication. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.

Submit your manuscript here: https://www.dovepress.com/journal-of-hepatocellular-carcinoma-journal