A Novel Four-Gene Signature Based on Nonsense-Mediated RNA Decay for Predicting Prognosis in Hepatocellular Carcinoma: Bioinformatics Analysis and Functional Validation

Jiaxin Zhao, Cheng Wang, Liang Zhao, Huiying Zhou, Rui Wu, Tao Zhang, Jiawei Ding, Junjie Zhou, Huilin Zheng, Lei Zhang, Tianci Kong, Jie Zhou, Zhenhua Hu

1Department of Hepatobiliary and Pancreatic Surgery, Department of Surgery, Fourth Affiliated Hospital, School of Medicine, Zhejiang University, Yiwu, Zhejiang Province, People’s Republic of China; 2Department of Radiology, Fourth Affiliated Hospital, School of Medicine, Zhejiang University, Yiwu, Zhejiang Province, People’s Republic of China; 3Zhejiang Provincial Collaborative Innovation Center of Agricultural Biological Resource Biochemical Manufacturing, School of Biological and Chemical Engineering, Zhejiang University of Science and Technology, Hangzhou, Zhejiang Province, People’s Republic of China; 4Department of Hepatobiliary and Pancreatic Surgery, Department of Surgery, First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang Province, People’s Republic of China

Correspondence: Zhenhua Hu, Department of Hepatobiliary and Pancreatic Surgery, Department of Surgery, Fourth Affiliated Hospital, School of Medicine, Zhejiang University, Yiwu, Zhejiang Province, People’s Republic of China, Tel/Fax +86-570-89935878, Email huzych@zju.edu.cn; Jie Zhou, Department of Hepatobiliary and Pancreatic Surgery, Department of Surgery, First Affiliated Hospital, School of Medicine, Zhejiang University, Hangzhou, Zhejiang Province, People’s Republic of China, Tel +86-13429661721, Email jzhou329@zju.edu.cn

Purpose: Nonsense-mediated RNA decay (NMD), a surveillance pathway for selective degradation of aberrant mRNAs, is associated with cancer progression. Its potential as a predictor for aggressive hepatocellular carcinoma (HCC) is unclear. Here, we present an innovative NMD risk model for predicting HCC prognosis.

Methods: The Cancer Genome Atlas (TCGA) data of 374 liver HCC (LIHC) and 50 normal liver samples were extracted. A risk model based on NMD-related genes was developed through least absolute shrinkage and selection operator Cox (LASSO-Cox) regression of the LIHC-TCGA data. Prognostic validation was done using GSE54236, GSE116174, and GSE76427 data. Univariate and multivariate Cox regression analyses were conducted to assess the prognostic value of the model. We also constructed nomograms for survival prediction. Tumor immune infiltration was evaluated using the CIBERSORT algorithm, and the tumor cell phenotype was assessed. Finally, mouse experiments verified UPF3B knockdown effects on HCC tumor characteristics.

Results: We developed a risk model based on four NMD-related genes (PABPC1, RPL8, SMG5, and UPF3B) and validated it using GSE54236, GSE116174, and GSE76427 data. The model effectively distinguished high- and low-risk groups corresponding to unfavorable and favorable HCC outcomes. Its prognostic prediction accuracy was confirmed through time-dependent ROC analysis, and clinical-use nomograms with calibration curves were developed. Single-cell RNA sequencing results indicated significantly higher expression of SMG5 and UPF3B in tumor cells. Knockdown of SMG5 and UPF3B inhibited HCC cell proliferation, invasion, and migration, while affecting cell-cycle progression and apoptosis. In vivo, UPF3B knockdown delayed tumor growth and increased immune cell infiltration.

Conclusion: Our NMD-related gene-based risk model can help identify therapeutic targets and biomarkers for HCC. Additionally, it assists clinicians in predicting the prognosis of HCC patients.

Keywords: nonsense-mediated RNA decay, hepatocellular carcinoma, biomarker, LASSO, survival, tumor immunity

Introduction

Cancer is a major global public health problem, with primary liver cancer being the sixth most common cancer and the third leading cause of cancer death worldwide. Primary liver cancer also demonstrates the highest incidence in transitioning countries. Hepatocellular carcinoma (HCC) development is associated with chronic hepatitis B or...
C virus infection, as well as with nonviral factors, such as iron overload syndrome, alcohol consumption, tobacco use, oral contraceptive pills, and aflatoxin exposure.\textsuperscript{2–4} A combination of ultrasound, computed tomography, magnetic resonance imaging, and serum alpha-fetoprotein measurement is conventionally used for primary diagnosis and screening of HCC.\textsuperscript{5} Conventional HCC treatment comprises surgery, chemotherapy, radiotherapy, liver transplantation, targeted therapy, and immunotherapy. Moreover, traditional Chinese medicine prescriptions have been noted to inhibit HCC cell growth and invasion potently. Emerging therapeutic approaches such as gene therapy may have potential in the treatment of HCC; relevant experimental studies have reported significant reductions in HCC cell proliferation without affecting normal hepatocyte function.\textsuperscript{6,7} Nevertheless, the prognosis of advanced HCC remains poor, highlighting the ongoing challenges related to improving outcomes in patients with HCC. Thus, methods for early diagnosis and prognosis prediction of HCC warrant development.

Biomarkers for early diagnosis of HCC may influence the treatment options for liver cancer considerably.\textsuperscript{8} Thus, many studies have investigated early biomarkers for liver cancer and prognosis prediction, making this an emerging, prominent research area. HCC exhibits considerable heterogeneity in terms of molecular characteristics and clinical manifestations.\textsuperscript{9} Therefore, research on biomarkers for early diagnosis and prognosis of HCC, as well as those for potential therapeutic targets for HCC, is highly relevant to improving overall HCC prognosis. An ideal biomarker for HCC should assist clinicians in detecting the disease early, exhibit high sensitivity and specificity, and be noninvasive and easily accessible. Such a biomarker may enable early HCC diagnosis and facilitate effective HCC treatment strategy implementation.\textsuperscript{10} Consequently, validating new, reliable biomarkers for HCC may aid in identifying novel targets for diagnostic or prognostic prediction of the disease. Through improvement of their early HCC detection and prognosis prediction accuracy, these biomarkers may improve overall survival (OS) rates in patients with HCC.

In cells, genetic information is transmitted from DNA to proteins through tightly regulated transcription and translation processes. Maintaining fidelity of these processes is essential for genetic stability and proper function of the cells. As such, cells have evolved various surveillance pathways to ensure mRNA accuracy; one of these pathways is called nonsense-mediated RNA decay (NMD).\textsuperscript{11} NMD is a translation-dependent quality control mechanism that scrutinizes mRNAs for errors possibly resulting in the formation of features, such as premature termination codons (PTCs), that trigger NMD activation. NMD involves selective degradation of faulty mRNA transcripts, which leads to the prevention of aberrant protein production and maintenance of gene expression fidelity.\textsuperscript{12} Being an RNA monitoring pathway operating at the transcription–translation interface, NMD has been extensively investigated in various model systems in organisms, such as yeast, \textit{Caenorhabditis elegans}, \textit{Drosophila}, and mice, as well as human cells.\textsuperscript{13} Recent studies have highlighted the significant effects of NMD dysregulation on cancer development and progression. NMD selectively downregulates tumor-suppressor gene (TSG) expression and maintains oncogene integrity, resulting in an imbalance in gene expression, which eventually promotes tumorigenesis and cancer progression.\textsuperscript{14} NMD has a regulatory role in the progression of genetic disorders, such as vascular hemophilia and retinitis pigmentosa. In these conditions, the deterioration of the pathway involved in NMD can exacerbate the conditions’ phenotype by triggering the expression of NMD-sensitive transcripts encoding functional proteins.\textsuperscript{15,16} Notably, several investigations have elucidated the pivotal role of NMD in the pathogenesis of various malignancies, including breast cancer and familial adenomatous polyposis. Furthermore, mutations in major TSGs such as \textit{BRCA1}, \textit{BRCA2}, \textit{TP53}, \textit{RB203}, and \textit{WT1} can potentially induce NMD activation, contributing to malignant tumor development.\textsuperscript{17,18} Impairments in NMD functionality, primarily attributable to decreased expression of the crucial protein UPF1, contribute to dysregulation of \textit{SMAD7} expression and aberrant mRNA surveillance mechanisms in HCC.\textsuperscript{19} Through upregulation of the UPF1/2 and SMG1/6/7 expression, NMD plays a crucial role in regulating microsatellite instability (MSI), a colorectal cancer initiation and progression biomarker.\textsuperscript{20} Some drugs may inhibit the progression and improve the prognosis of cancer by targeting the NMD pathway; these drugs include homoharringtonine, NMDI-1, as well as small-molecule drugs, which are already in use for cancer treatment or under investigation in clinical trials.\textsuperscript{21–23} However, the NMD mechanisms underlying the development and prognosis of HCC remain unclear. Herein, we, for the first time, propose an NMD risk model, for elucidating the prognostic significance of NMD in HCC patients. Our results may improve the current understanding of the roles of NMD-mediated gene regulation in HCC, as well as cancer in general, and facilitate the exploration of potential therapeutic interventions.
Materials and Methods

Data Access and Normalization

We extracted and used the RNA sequencing (RNA-seq) data of 50 normal liver tissue and 374 liver HCC (LIHC) samples from The Cancer Genome Atlas (TCGA) database (https://portal.gdc.cancer.gov/) for gene expression analysis. We also used gene expression data in the clinical data files GSE54236, GSE116174, and GSE76427 from the Gene Expression Omnibus (GEO) database (https://www.ncbi.nlm.nih.gov/geo/). To ensure data comparability and remove potential batch effects, we performed batch normalization by using the R package “sva” (version 4.2.1). This step enabled adjustments for all systematic differences between the datasets and enhancement of the reliability of our subsequent analyses. By applying batch normalization, we could minimize unwanted variations due to technical factors and focus on biological variations of interest.

NMD-Related Gene Identification

We retrieved NMD-related gene data from the Gene Set Enrichment Analysis (GSEA) database. Moreover, we performed a comprehensive differential expression analysis of TCGA data on the NMD-related genes in 50 normal liver and 374 LIHC samples by using the R package “DESeq2.” Here, 15 DEGs, with statistical significance (P < 0.05) and a minimum optimal fold change (FC) [|log (FC)| ≥ 0.585], were identified. Next, we used the R packages “ComplexHeatmap” and “ggplot2” to generate heatmaps and volcano plots, respectively. To explore potential protein-protein interactions (PPIs) among the identified NMD-related genes, we used data from the STRING database (https://cn.string-db.org/). We set a minimum required interaction score of 0.4 (ie, medium confidence) to ensure the reliability of our PPI analysis. Finally, alterations in the selected genes in liver cancer data were analyzed using cBioPortal (https://www.cbioportal.org/).

HCC Prognosis–Related Gene Identification

We performed univariate Cox regression analysis by using the R package “survival.” The significance of the relationship between NMD-related gene expression and OS in patients with HCC was assessed, with P < 0.05 considered the screening threshold. Next, we applied least absolute shrinkage and selection operator (LASSO) Cox regression. LASSO Cox regression is typically used for survival analysis, particularly to identify variables with a significant impact on survival outcomes. Specifically, the Cox proportional hazards model is used for assessing the effects of variables on survival time or time to an event, and the LASSO is employed for selecting features and regularizing the model. Because LASSO Cox regression facilitates the identification of crucial prognostic variables and the construction of corresponding risk models, it has been extensively applied in clinical and biological research, particularly in cancer studies, for discovering potential prognostic biomarkers and building prediction models. We used this method to account for gene covariance and reduce the number of modeled genes, facilitating the identification of a subset of NMD-related genes most strongly associated with OS in the LIHC samples. Finally, we performed multivariate Cox regression analysis based on the results of univariate Cox regression.

Construction and Validation of Our Risk Model Based on NMD-Related Genes

We calculated the risk scores by using centralized and standardized mRNA expression data from the LIHC training set. The patients were stratified into high- and low-risk groups based on the median risk scores in this study. To gain insights into the relationships between variables and correlation and risk heatmaps were generated using the R packages “Psych” and “pheatmap.” To validate our risk model’s prognostic value for HCC, we plotted Kaplan–Meier (K-M) curves by using the survival analysis functions of the R package “survival.” These curves provided valuable information regarding the survival rates of the different LIHC risk groups. Receiver operating characteristic (ROC) curves were used to assess the predictive value of our risk models regarding HCC prognosis. Principal component analysis (PCA) and t-distributed stochastic neighbor embedding (tSNE) were used to evaluate our risk model’s validity. Furthermore, univariate and multivariate Cox regression analyses were performed to determine the independent prognostic value of our risk model. Nomogram plots were generated using the R package “rms”, specifically for TCGA dataset. Finally, we used calibration curves to assess the nomogram plots’ accuracy.
GSEA and Single-Cell RNA-Seq Analysis

We performed GSEA to identify the biological processes and signaling pathways differing between the high- and low-risk groups in our TCGA-LIHC cohort, with the C2 (c2.cp.kegg.v7.5.1.symbols.gmt) and C5 (c5.go.v2023.2.Hs.symbols.gmt) subcollections as our reference gene sets. To assess the statistical significance of our findings, we performed 1000 permutation analyses and established a threshold for significance at \( P < 0.05 \) (NOM-p < 0.05, FDR-q < 0.05, FWER-p < 0.05). We analyzed SMG5 and UPF3B expression in liver cancer cells using single-cell RNA-seq (scRNA-seq) data from the liver cancer database VGDS (http://vgds.bitlife.com.cn/), the first multidimensional data analysis platform integrating bulk RNA-seq, scRNA-seq, and literature data.

Analysis of Tumor Mutation Burden and Immune Characteristics

The tumor microenvironment plays a significant role in the behavior of neoplastic cells and their response to drug treatments. Tumor mutation burden (TMB) is a factor affecting the immunological response of cancer cells. Here, we used the CIBERSORT algorithm to compute the contents of immune cell subpopulations in TCGA-LIHC samples.

Single-sample GSEA was then used to compare the variations in the range of immune cell and function types. We evaluated the mutation frequency of patients with HCC in the two risk groups using the R package “maftools.” The top 20 genes with the highest mutation frequency were selected, and a waterfall diagram was plotted to describe the mutation. We also combined the mutation data with the survival data to obtain the relationship between our risk model, transformation, and survival.

Cell Culture

We procured the human HCC cell lines MAHLAVU and Huh-7 from the American Type Culture Collection (Manassas, VA, USA). The cells were cultured in Dulbecco’s modified Eagle medium (DMEM; Gibco, USA) supplemented with 10% fetal bovine serum (Procell, USA) and 1% penicillin–streptomycin solution (containing 10,000 U/mL of penicillin and 10 mg/mL streptomycin; VivaCell) at 37 °C under 5% CO₂ in a humidified atmosphere.

Cell Transfection

We seeded the cultured cells in 6-well plates at 2×10⁵ cells per well and incubated at 37 °C under 5% CO₂ in a humidified atmosphere for 24 h until 70% confluency was reached. Subsequently, complete DMEM was replaced with serum-free DMEM, and siRNA transfection was performed using Lipofectamine 3000, according to the manufacturer’s instructions. The siRNAs used for UPF3B and SMG5 knockdown were as follows:

- si-UPF3B-1: 5’-GGACAAAUUAAGGAUGAATT-3’
- si-UPF3B-2: 5’-GGUCAAGAGAGAUCGAAUATT-3’
- si-SMG5-1: 5’-GCCGCUUCAUUGUCAUCAUTT-3’
- si-SMG5-2: 5’-GGCCUUUGAUGCUGAAACATT-3’

We also used si-NC for transfection control.

Western Blotting

To investigate the potential influence of UPF3B and SMG5 on cellular protein expression, we lysed cells treated with si-NC and siRNAs against the genes by incubating them with a 100:1 mixture of RIPA buffer (Beyotime, China) and phenylmethylsulfonyl fluoride (EpiZyme, China) for 5 min on ice. The cells were scraped off using a cell scraper, and the protein samples were then diluted with a loading buffer (Beyotime, China). The proteins were denatured by heating the extracted proteins at 95 °C for 15 min and cooling them rapidly on ice before use. Subsequently, the proteins were separated through sodium dodecyl sulfate polyacrylamide gel electrophoresis and transferred onto polyvinylidene fluoride membranes. The membrane was blocked for 3 h and then incubated with relevant primary antibodies (Abcam,
UK) at 4 °C overnight. Finally, to detect and identify the protein bands, we incubated the membranes with a relevant secondary antibody (Proteintech, China) for 2 h.

**Cell Counting Kit-8 and Colony Formation Assays**

To investigate the effects of gene regulation on cell proliferation, we conducted Cell Counting Kit-8 (CCK-8) assay on Huh-7 cells treated with si-NC and siRNAs against the target genes. In brief, the cells were seeded in 96-well plates at a density of 3,000 cells per well, with three replicates per group. At 0, 24, 48, and 72 h, 10 µL of CCK-8 reagent (Beyotime, China) was added to each well, followed by incubation at 37 °C for 2 h. Next, cell viability was assessed on the basis of absorbance at 450 nm.

Huh-7 cells were also seeded in a six-well plate at 500 cells per well in droplets. Next, 3 mL of medium was added to each well, and the plate was monitored regularly. By day 12, the plate demonstrated cell colony formation. The colonies were fixed using methanol for 30 min and stained with crystal violet (Beyotime).

**Wound Healing Scratch Test and Transwell Assay**

To evaluate the migration capability of our HCC cells, the cells were seeded onto a 12-well plate and incubated for 24 h. Next, a scratch was made in the cell monolayer using a 10-µL pipette tip, and the debris generated during scratching was removed using phosphate-buffered saline. Subsequently, 1 mL of serum-free DMEM was added to each well, and images of the gaps created were captured under a microscope. Cell migration was observed at 0 and 48 h after scratch creation, and the analysis of cell migration was performed using Image J.

Moreover, approximately 50,000 cells from each group were seeded into the upper chamber of transwell plates (Corning, USA), and 500 µL of complete DMEM was added to the lower chamber of the plates, followed by incubation for 48 h. Next, the cells were fixed with methanol for 30 min and then stained with crystal violet for 20 min.

**Nude Mouse Model of Tumor Implantation**

Our animal experiment protocol was evaluated by the Ethical Review Committee for Laboratory Animal Welfare of Zhejiang University (approval number: ZJU20230416), following the guidelines of “Guidelines for Ethical Review of Animal Welfare (GBT 35892–2018)”.

Male C57BL/6 nude mice were purchased from Jicuiyaokang Company (Jiangsu, China) and treated using an experimental method reported previously.

In brief, each mouse was injected with 2×10⁶ Hepa1-6 cells (n = 3 mice/group). After 5 days, 5 nmol of si-UPF3B (in 50 µL of 0.9% NaCl) was injected into the cancerous tissues of each mouse once every 3 days for 2 weeks. After approximately 2 weeks, all mice were killed, and their tumors were resected and weighed.

**Flow Cytometry Assay**

A cell-cycle staining kit (MULTI SCIENCES) was employed for our flow cytometry analysis. To prepare the samples, 1×10⁶ cells were mixed with 1 mL of a DNA staining solution and 10 µL of a permeabilization solution. The mixture was thoroughly mixed through vortexing for 5–10 s. This was followed by incubation at room temperature for 30 min in the dark. Cell apoptosis was assessed through Annexin V-FITC/PI dual staining, followed by flow cytometry analysis on the BD FACSCanto II System (BD, USA).

The method used here to prepare single-cell suspensions from mouse tumors is described elsewhere. In brief, mice were sacrificed 14 days after the treatments; the tumors were cut into pieces, digested, and suspended. Next, the suspensions were filtered using a 70-µm nylon mesh filter. Next, the Zombie Aqua Fixable Viability Kit was used for cell membrane staining and then incubated with flow cytometry antibodies against CD45 and CD3. Finally, the sample was filtered again through a 40-µm nylon filter to obtain single-cell suspensions. These suspensions were assessed through flow cytometry on the BD FACSCanto II System.

**Statistical Analysis**

All statistical analyses were performed using R (version 4.2.1) and GraphPad Prism (version 8). Between-group differences were assessed using the t-test and the Mann–Whitney U-test. Survival analyses were performed using...
K-M curves, Log rank tests, and Cox proportional hazards regression models. Correlations were assessed using the Spearman correlation test. Cell-cycle outcomes were analyzed using Flow Jo (version 10). Each experiment was repeated three times. In general, a $P$ value of $<$0.05 was considered to indicate statistical significance; moreover, *** and **** were used to denote statistical significance at $P$ values of $<$0.05, $<$0.01, and $<$0.001, respectively.

Results

NMD-Related Gene Screening

Figure 1 illustrates the flow of this study, aimed at identifying a novel NMD-based gene signature for predicting HCC prognosis. We collected 424 data samples from TCGA database (374 LIHC and 50 normal liver tissues). We also obtained 117 NMD-related genes from the GSEA dataset and identified 15 DEGs using the DESeq2 package, with $|\text{Log (FC)}| \geq 0.585$ and $P < 0.05$ as thresholds (Figure 2A and B). Figure S1 presents the Gene Ontology (GO) terms and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enriched in the 15 DEGs. The PPI network of these 15 genes is illustrated in Figure 2C.

Through univariate Cox regression analysis, we identified five genes associated with HCC prognosis (Figure 2D). Figure 2E presents the interrelationship among these five genes. Furthermore, LIHC samples (Figure 2F) were noted to express these five genes predominantly in their amplified forms.

Construction of Our Risk Model Based on NMD-Related Genes

To investigate the prognostic value of NMD-related genes in patients with HCC, we conducted comprehensive transcriptomic and clinical data analyses of TCGA data through LASSO Cox regression. Consequently, we selected the four most optimal candidate genes and included them in the final risk model (Figure 3A and B and Table S1):
Risk Score = (0.0662 × PABPC1 expression) + (0.0811 × RPL8 expression) + (0.3401 × SMG5 expression) + (0.3232 × UPF3B expression)

The expression of all four genes demonstrated predictive value for HCC prognosis (Figure S2A–D). Next, we divided the LIHC samples into a high-risk group, comprising those with risk scores higher than the median, and a low-risk group, comprising those with risk scores lower than the median. To evaluate our risk model’s performance, we performed PCA and tSNE on TGCA-LIHC data (Figure 3C and D). The results demonstrated that our model effectively distinguished the LIHC samples based on their risk levels. Thus, our risk model may accurately stratify patients with HCC based on their genetic risk profiles.
To evaluate the prognostic value of our HCC risk model, we performed K-M survival analysis on the high- and low-risk groups within TCGA-LIHC data. Notably, the results revealed a significant difference in OS between the two groups: it was significantly lower in the high-risk group than in the low-risk group (Figure 4A). The risk curves and scatter plots further demonstrated a positive correlation between risk scores and mortality rates in patients with HCC (Figure 4C and E). Additionally, the ROC curve analysis confirmed the accuracy and reliability of our risk model in predicting HCC prognosis (Figure 4G).

Next, we performed prognostic validation using data from the GEO datasets GSE54236, GSE116174, and GSE76427. We first assessed data distribution before and after normalization through PCA, and the results demonstrated that the data were comparable after normalization (Figures S3 and S4). For these dataset data, the K-M survival analysis, risk curves, and scatter plots also indicated a positive correlation between risk scores and mortality in patients with HCC (Figure 4B, D and F). Moreover, the ROC curves substantiated the considerable predictive ability of the risk scores (Figure 4H). Therefore, our risk model demonstrated a robust prognostic potential, suggesting its utility in predicting HCC prognosis accurately.
Figure 4 Analysis of risk model and survival. (A) Survival curve for TCGA-LIHC data in high- and low-risk groups. (B) Survival curve for GEO-LIHC data in high- and low-risk groups. (C and D) Distribution of risk scores and the risk profile. (E and F) Distribution of risk scores and the survival status. (G and H) ROC curve analysis.
Independent Prognostic Analysis of Our Risk Model and Clinical Characteristics

We performed a clinical correlation analysis to assess our risk model’s clinical relevance further. Table S2 lists the clinical data of high- and low-risk group patients. Our findings revealed no significant differences in the risk score based on gender (P = 0.65) and age (P = 0.92), indicating that our risk model was not influenced by gender or age (Figure 5A and 5B). Therefore, we did not consider age and gender as confounders in our risk model based on NMD-related genes. In contrast, the risk score was strongly and positively correlated with the tumor stage; in other words, the higher the risk score, the more advanced may be the tumor stage (Figure 5C). Notably, no significant differences were noted for stage IV HCC, possibly because of the small sample size, which led to statistical errors.

Next, univariate and multivariate Cox regression analyses were performed to evaluate the prognostic value of our risk model. In the univariate Cox regression analysis, the hazard ratio [HR; 95% confidence interval (CI)] for the risk score was 2.978 (2.034–4.360); therefore, a higher risk score may be associated with a significantly higher risk of poor prognosis (P < 0.001; Figure 5D). In the multivariate analysis, the HR (95% CI) for the risk score was 2.584 (1.765–3.782), which further confirmed the independent prognostic value of the risk scores (P < 0.001; Figure 5E).

To facilitate the application of our risk model in clinical practice, we developed a nomogram combining risk scores and clinicopathologic parameters to predict the survival rates of patients with HCC at 1, 3, and 5 years (Figure 5F). The calibration curves demonstrated that the nomogram exhibited good agreement between predicted and actual survival rates, indicating the high predictive ability of the nomogram (Figure 5G). Then, we correlated NMD scores with sorafenib IC50 values and found a significant association. Patients with high NMD expression had lower sorafenib IC50 scores, indicating the high predictive ability of the nomogram (Figure 5H).

Dovepress

Taken together, the aforementioned results collectively highlighted our model’s clinical utility and predictive accuracy for HCC. These findings have major implications for personalized treatment strategies and outcome prediction in patients with HCC.

GSEA and scRNA-Seq Analysis for NMD-Related Genes in High-Risk HCC Patients

To investigate the functional enrichment of NMD-related genes in high-risk HCC patients, we conducted GSEA. As shown in Figure 6A, these were the top 10 enriched KEGG pathways in the high-risk HCC patients: basal transcription factors, cell cycle, neurotrophin signaling pathway, oocyte meiosis, progesterone-mediated oocyte maturation, pyrimidine metabolism, regulation of autophagy, RNA degradation, spliceosome, and ubiquitin-mediated proteolysis. Moreover, as shown in Figure 6B, these were the top 10 enriched GO terms pathways in the high-risk HCC patients: DNA-templated transcription initiation, mRNA catabolic process, negative regulation of amide metabolic process, positive regulation of mRNA metabolic process, regulation of mRNA catabolic process, regulation of mRNA metabolic process, RNA catabolic process, RNA destabilization, mRNA 3’ untranslated region (UTR) binding, and translation regulator activity. We also performed scRNA analysis on SMG5 and UPF3B expression; as shown in Figure 6C and D, HCC cells, particularly in those from patients with metastasis, demonstrated significant overexpression of SMG5 and UPF3B.

Immune Infiltration and TMB in High-Risk HCC Patients

The TME is critical for cancer progression, particularly for tumor immunity. We further investigated the association between immune infiltration and NMD. As shown in Figure 7A, the risk score was significantly and negatively correlated with mast cells, neutrophils, natural killer (NK) cells, and helper and regulatory T cells. Moreover, the risk score demonstrated a significant negative correlation with chemokine receptors, cytolytic activity, and type I and II interferon responses (Figure 7B). These findings collectively suggested that high-risk HCC patients exhibit an immune profile favoring tumor deterioration.

We also investigated the relationship between TMB and the risk score. The results demonstrated that as the risk score increased, the TMB also increased (Figure 7C). Notably, the high-risk group displayed a significantly higher TP53 mutation rate (Figure 7D and E). Subsequently, we evaluated patient survival by combining TMB and the risk score. The results revealed that patients with a high mutation rate had a poor prognosis (Figure 7F); moreover, within the high-risk group, a high mutation rate was associated with the worst survival outcomes (Figure 7G).
Effects of SMG5 and UPF3B Knockdown on HCC Cell Proliferation, Migration, and Invasion

To explore the role of NMD in HCC cells, we targeted SMG5 and UPF3B, two of the four pivotal genes involved in NMD. First, we knocked down the expression of SMG5 and UPF3B in the HCC cell lines MAHLAVU and Huh7 and
verified the knockdown efficiency through Western blotting (Figure 8A). Next, we investigated the influence of SMG5 and UPF3B on HCC cell proliferation by using CCK-8 assays. Our findings revealed that at 24, 48, and 72 h, SMG5 and UPF3B knockdown significantly reduced proliferation in HCC cells compared with the si-NC-treated cells group (Figure 8B). Furthermore, HCC cells with SMG5 and UPF3B knockdown demonstrated a substantial reduction in the clonogenic potential in the colony formation assay (Figure 8C), a considerable decline in the invasive ability in the transwell assay (Figure 8D and E), and a notable decrease in the migratory ability the wound healing scratch test (Figure 8F–I).

Moreover, the GSEA results demonstrated that the mTOR pathway met the following criteria: NOM-p < 0.05, FDR-q < 0.05, and FWER-p < 0.05. Consequently, we validated the pathway by focusing on UPF3B, which demonstrated the most significant phenotype in our experiments. The results demonstrated that UPF3B knockdown significantly inhibited the mTOR pathway, suggesting that UPF3B has a tumor-promoting role via the mTOR pathway (Figure 8J).

Taken together, these results indicate the potential involvement of NMD-related genes such as SMG5 and UPF3B in the regulation of HCC cell proliferation, clonogenicity, invasion, and migration and confirm the role of key NMD-related genes in HCC progression.
Effects of SMG5 and UPF3B Knockdown on Apoptosis and Cell Cycle in HCC Cells

To further elucidate the function of essential NMD-related genes in HCC progression, we performed flow cytometry and assessed the effects of SMG5 and UPF3B knockdown on cell cycle and apoptosis in HCC cells. Our results revealed that SMG5 and UPF3B knockdown significantly prolonged the G2 phase of the cell cycle, indicating that the growth of the HCC cells was blocked at the G2 phase, further delaying cell proliferation (Figure 9A and B). Moreover, our cell apoptosis analysis demonstrated that SMG5 and UPF3B knockdown considerably increased the apoptosis rate in HCC cells (Figure 9C and D). Therefore, through cell-cycle inhibition and apoptosis pathway alteration, SMG5 and UPF3B
potentially underlie the expression of various phenotypes in HCC cells. These results provide further insights into the mechanisms underlying NMD’s involvement in HCC pathogenesis and may aid in developing more targeted therapies for the devastating disease.

Effects of UPF3B Knockdown on Tumor Growth and Immune Infiltration

To evaluate the potential anticancer immune effects of UPF3B knockdown on the HCC cell line Hepa1-6, we subcutaneously implanted the cells in the right flank of C57BL/6 nude mice. After 5 days, the mice were randomly...
divided into two groups. After the tumor diameter reached 6–8 mm, the treatment group received an initial injection of 5 nM si-UPF3B (in 50 μL of 0.9% NaCl) every 3 days over 2 weeks. Next, the tumors were excised, weighed, and photographed; the results are presented in Figure 9E. Subsequently, single-cell suspensions were obtained from the tumors and subjected to flow cytometry for viable tumor-infiltrating CD45^+ and CD45^+CD3^+ immune cells; the results revealed a significant elevation in the number of CD45^+ immune cells and CD45^+CD3^+ immune cells in the tumors of the **UPF3B** knockdown group (Figure 9F).

**Discussion**

HCC, which represents a prevailing form of liver cancer, is associated with a substantial global health burden. Despite recent scientific advancements and resulting increases in patient survival rates, the efficacy of liver cancer treatments remains low. As such, the need to investigate molecular mechanisms underlying HCC progression is increasing. Aberrant gene expression is a major factor involved in HCC progression; this highlights the significance of exploring the associated molecular pathways. NMD is a conserved RNA degradation pathway, in which RNAs containing truncating mutations, such as nonsense, frameshift, and specific splice-site mutations, are targeted selectively and their translation is terminated prematurely. Recent studies have emphasized the vital role of NMD in cancer progression. For instance, mutations in **UPF1**, a crucial gene involved in NMD, are significantly associated with pancreatic squamous carcinoma prognosis. Moreover, NMD inhibition can reduce tumorigenicity in human fibrosarcoma cells. Therefore, in the present study, we established a novel HCC risk model based on the four major NMD-related genes **PABPC1**, **RPL8**, **SMG5**, and **UPF3B**, which can aid in predicting prognosis and guiding treatment in patients with HCC. In particular, we used LASSO-Cox regression analysis to construct our risk model, and it demonstrated a substantial predictive value for HCC prognosis. We then combined our risk model and clinical parameters to construct a nomogram, which also demonstrated high predictive value. Our model is the first-ever risk model constructed for a liver disease based on NMD-related genes.

**PABPC1** promotes cell proliferation and metastasis by regulating **COL12A1** expression in pancreatic adenocarcinoma. Furthermore, **PABPC1** expression is downregulated in several malignant gliomas, and therefore, it may be a predictor of preneural subtypes of gliomas; notably, in patients with glioma, high **PABPC1** expression is significantly associated with a favorable prognosis. In contrast, our risk model indicated that increased **PABPC1** expression is associated with poor HCC prognosis. **RPL8** is a potential candidate gene associated with poor prognosis and immune infiltration in HCC. **RPL8**, a ribosomal subunit protein, plays essential roles in protein synthesis and cell growth regulation.

**SMG5** and **UPF3B** are key genes related to NMD. In the NMD pathway, **SMG5** plays a crucial role by participating in **SMG1C** complex formation and interacting with other NMD factors, which eventually leads to degradation of aberrant mRNAs. In particular, by regulating **SMG6** endonuclease activity, **SMG5** facilitates PTC-containing mRNA recognition and degradation, thus maintaining gene expression integrity. Therefore, in the NMD pathway, **SMG5** is essential for its contribution to cell quality control and gene expression. **SMG5** may be differentially expressed in several cancers, and it may be significantly correlated with low-grade glioma prognosis, immune microenvironment, and immune checkpoints. **SMG5** promotes **UPF1** dephosphorylation by recruiting protein phosphatase 2A, and the phosphorylation–dephosphorylation cycle is essential for NMD activity. Independent of its association with the exon junction complex, **UPF3B** facilitates NMD activation. Furthermore, targeted inhibition of **UPF3B** expression could be a potential therapeutic strategy for diseases caused by nonsense mutations, thereby restoring regular gene expression and protein function. **UPF3B** expression is significantly correlated with the pathological staging of several cancer types, including esophageal carcinoma, renal chromophobe, HCC, and skin cutaneous melanoma. Furthermore, in many cancer types, higher **UPF3B** transcript levels are associated with significantly poorer prognoses. However, no study has reported the crucial roles of **SMG5** and **UPF3B** in hepatocarcinogenesis and progression by using a risk model based on NMD-related genes.

Next, we confirmed the significant proliferative, migratory, and invasive roles of **SMG5** and **UPF3B** in HCC cells through in vitro experiments. The results also confirmed that these genes promote HCC, indicating that they are regulated by NMD and are integral to shaping the TME conducive to HCC cell growth; in other words, their expression is
Figure 9  Tumor cell cycle and apoptosis after SMG5 and UPF3B knockdown. (A and B) Changes in the cell cycle of Huh7 and MAHLAVU cells after si-SMG5 and si-UPF3B treatment. (C and D) Changes in apoptosis in Huh7 and MAHLAVU cells after si-SMG5 and si-UPF3B treatment. (E) Tumor weights and representative photographs of resected tumors at sacrifice (n = 3 per group). (F) Percentage of viable tumor-infiltrating CD45+ immune cells and T cells (CD45+CD3+). *P < 0.05, **P < 0.01, ***P < 0.001.
associated with poor HCC prognosis. Furthermore, we assessed the cell cycle analysis and apoptosis in HCC cells by exploring the pathways through which SMG5 and UPF3B promote HCC proliferation. The results demonstrated that SMG5 and UPF3B promote HCC cell proliferation through G2 cell-cycle arrest induction and apoptosis inhibition.

The mTOR pathway is overexpressed in nearly 50% of patients with HCC, and the dysregulated activation of this pathway affects a wide range of processes, including cell proliferation, metabolism, tumor cell differentiation, lipid metabolism, autophagy, and epithelial–mesenchymal transition.\(^{42}\) We noted that UPF3B knockdown significantly inhibited the mTOR pathway; in other words, our result, for the first time, indicated that UPF3B may influence HCC progression by modulating the mTOR pathway. In general, our novel insights into the roles of the key NMD-related genes SMG5 and UPF3B in HCC development and progression may facilitate the development of targeted therapies for HCC.

Our GSEA results revealed that the high-risk HCC group had significant enrichments in the pathways related to basal transcription factors, cell cycle, RNA degradation, and spliceosome, as well as in the processes related to regulation of mRNA catabolic process, RNA destabilization, mRNA 3′-UTR binding, and translation regulator activity. Moreover, our scRNA-seq analysis results indicated that NMD-related gene expression was significantly higher in patients with highly malignant tumors than in other patients. These results confirmed that mRNA degradation in the high-risk HCC group is regulated significantly differently from that in the low-risk HCC group.

In tumors, NMD inhibition can enhance tumor immunity. Tumor cells can evade the immune system mainly because they do not express the potent tumor rejection antigens. The antigenicity of disseminated tumors may be increased through NMD inhibition; this may lead to the induction of new concentrated antigenic expression on tumor cells and eventually to immune recognition and rejection.\(^{43}\) A study reduced NMD activity in the breast cancer cells of mice through knockdown of UPF2, a critical factor involved in NMD; the results demonstrated that reducing NMD activity led to production of novel mRNA isoforms and NMD-sensitive transcripts, which encode tumor neoantigens in breast cancer cells.\(^{44}\) Moreover, studies have explored whether NMD affects tumor inflammation by inhibiting the NMD pathway in tumor cells; they noted tumor inflammation induction, which improved the cells’ responses to immune checkpoint blockade.\(^{45}\) Cancer cells can use the NMD pathway to downregulate essential TSGs and adapt to an aggressive immune microenvironment by adjusting the NMD activity.\(^{46}\) However, NMD plays two roles in tumor development: (1) Cancer cells use NMD to reduce TSG expression by selecting for mutations leading to transcriptome disruption. (2) Cancer cells adapt to changes in the TME by adjusting NMD activity.

In the current study, the high-risk group demonstrated poorer levels of cancer immunity: the numbers of mast cells, neutrophils, NK cells, and helper and regulatory T cells were relatively low, and the responses to cytotoxicity and interferon were relatively poor. Patients with lower cytotoxicity responses demonstrated lower 5-year survival rates than those with higher cytotoxicity responses.\(^{47}\) Moreover, type I interferons can block cell-cycle progression, induce apoptosis, facilitate clearance, and prevent metastasis.\(^{48}\) Therefore, in our high-risk group, the dysregulated mRNA decay created an immune microenvironment more suitable for cancer progression, which led to poor HCC prognosis. Furthermore, the results of our in vivo experiments demonstrated that the immune infiltration levels in HCC tissues increased after UPF3B knockdown, further confirming the close association between NMD and the TME.

NMD frequently leads to TSG mutations that promote cancer development. In other words, TSGs represent the most pertinent gene class associated with NMD. Inactivating the function of TSGs such as TP53, RB1, or PTEN can lead to unrestrained tumor growth via various mechanisms (eg, overriding cell-cycle control). This is because TSGs act as “brakes” in tumorigenesis, and loss of their function can contribute to cancer development.\(^{49,50}\) In the current study, we observed a strong correlation between NMD and TP53 mutations in the high-risk HCC groups. TP53, a well-known TSG, plays a crucial role in regulating cell-cycle control and preventing cancer development. Our results indicated that an increase in the expression of NMD-related genes may lead to an increase in the TP53 mutation rate, potentially leading to cell-cycle control override and tumor growth promotion. To explore its prognostic significance further, we integrated our risk model based on NMD-related genes with tumor mutation load and survival analysis. Notably, the high-risk and high-mutation groups demonstrated the most unfavorable prognosis. Therefore, a combination of NMD-based high-risk status with increased TMB may be a valuable prognostic indicator in HCC. Our findings enhance the
current understanding of the molecular mechanisms underlying cancer development and may have implications for the identification of potential prognostic markers and therapeutic targets for HCC.

Our results offered valuable insights into the roles of SMG5 and UPF3B in HCC cells in vitro and indicated that UPF3B, a critical NMD-related gene, can enhance tumor immune infiltration in vivo. However, this study has a limitation: we did not characterize the phenotypes of immune cells. Future studies should, therefore, focus on elucidating immune cell phenotypes specific to NMD in HCC; their results may enhance the current understanding of NMD’s role in the TME and systemic interactions in HCC.

In conclusion, we reported a novel risk model based on NMD-related genes for the prediction of HCC prognosis. Furthermore, through functional enrichment analysis and in vivo experiments, we have elucidated the promoting role of crucial NMD-related genes in HCC progression and their impact on HCC immunity. Our results enhance the current understanding of the promotional role of NMD in HCC, potentially enabling the development of strategies for improved prognosis prediction and personalized treatment of patients with cancer in the future.

Data Sharing Statement
The datasets used in this study can be found in the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases. The data for this study are available in the Cancer Genome Atlas (https://portal.gdc.cancer.gov/repository?facetTab=files); GSE54236, GSE116174 and GSE76427 datasets are available from the following website (https://www.ncbi.nlm.nih.gov/).

Ethics Approval and Consent to Participate
This study involved in the sample information and clinical data are from the TCGA database (https://portal.gdc.cancer.gov/) and the GEO database (https://www.ncbi.nlm.nih.gov/geo/); the liver cancer cells were purchased from Wuhan Pricella Biotechnology Co., Ltd. The above study meets the conditions for exemption from ethical review in the Measures for Ethical Review of Life Sciences and Medical Research Involving Human Beings in China, so we obtained an ethical review exemption from the Human Research Ethics Committee of the Fourth Affiliated Hospital of Zhejiang University School of Medicine. The animal study has been evaluated by the Laboratory Animal Welfare Ethics Review Committee of Zhejiang University.

Acknowledgments
We acknowledge TCGA, GEO and VGDS database for providing their meaningful data.

Author Contributions
All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding
This study was supported by National key research and development program (2022YFA1104600), Key Research and Development Program of Zhejiang Province (No.2020C03057), Key Science and Technology Program of Zhejiang Province (No. WKJ-ZJ-1818).

Disclosure
The authors declare no competing interests in this work.

References