



OPEN ACCESS

EDITED BY

Nan Zhang,
Harbin Medical University, China

REVIEWED BY

Wantao Wu,
Central South University, China
Siqing Sun,
The First Affiliated Hospital of China
Medical University, China

*CORRESPONDENCE

Yan Wu
✉ cmuwuyan@163.com

RECEIVED 06 March 2023

ACCEPTED 20 April 2023

PUBLISHED 16 May 2023

CITATION

Bai Y, He J, Ma Y, Liang H, Li M and Wu Y
(2023) Identification of DNA repair gene
signature and potential molecular subtypes
in hepatocellular carcinoma.
Front. Oncol. 13:1180722.
doi: 10.3389/fonc.2023.1180722

COPYRIGHT

© 2023 Bai, He, Ma, Liang, Li and Wu. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that
the original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution or
reproduction is permitted which does not
comply with these terms.

Identification of DNA repair gene signature and potential molecular subtypes in hepatocellular carcinoma

Yi Bai¹, Jinyun He², Yanquan Ma¹, He Liang³,
Ming Li⁴ and Yan Wu^{5*}

¹Department of Critical Care Medicine, Panjin Liaoyou Baoshihua Hospital, Liaoning, China,

²Department of hepatobiliary surgery, Panjin Liaoyou Baoshihua Hospital, Liaoning, China,

³Department of integrated Chinese and Western medicine, Panjin Liaoyou Baoshihua Hospital, Liaoning, China, ⁴Fuxin Municipal Discipline Inspection Commission, Liaoning, China, ⁵Department of rheumatology and immunology, Panjin Liaoyou Baoshihua Hospital, Liaoning, China

DNA repair is a critical factor in tumor progression as it impacts tumor mutational burden, genome stability, PD-L1 expression, immunotherapy response, and tumor-infiltrating lymphocytes (TILs). In this study, we present a prognostic model for hepatocellular carcinoma (HCC) that utilizes genes related to the DNA damage response (DDR). Patients were stratified based on their risk score, and groups with lower risk scores demonstrated better survival rates compared to those with higher risk scores. The prognostic model's accuracy in predicting 1-, 3-, and 5-year survival rates for HCC patients was analyzed using receiver operator curve analysis (ROC). Results showed good accuracy in predicting survival rates. Additionally, we evaluated the prognostic model's potential as an independent factor for HCC prognosis, along with tumor stage. Furthermore, nomogram was employed to determine the overall survival year of patients with HCC based on this independent factor. Gene set enrichment analysis (GSEA) revealed that in the high-risk group, apoptosis, cell cycle, MAPK, mTOR, and WNT cascades were highly enriched. We used training and validation datasets to identify potential molecular subtypes of HCC based on the expression of DDR genes. The two subtypes differed in terms of checkpoint receptors for immunity and immune cell filtration capacity. Collectively, our study identified potential biomarkers of HCC prognosis, providing novel insights into the molecular mechanisms underlying HCC.

KEYWORDS

hepatocellular carcinoma, DNA repair gene, prognosis, molecular subtypes of HCC, immune infiltration

1 Introduction

According to the 2018 Global Cancer Statistics report, Hepatocellular Carcinoma (HCC) ranks 6th among malignancies and is the 4th leading cause of cancer-related mortality (1, 2). Despite significant advancements in HCC treatment, the outcomes are still unsatisfactory (3, 4). Therefore, identifying novel therapeutic targets and diagnostic biomarkers for HCC is crucial to improve patient prognosis. The DNA damage response (DDR) pathway is considered a potential source of therapeutic targets as damaged DNA is a hallmark of cancerous cells (5).

Research studies have reported that genes involved in DDR pathways, such as nucleotide and base excision and mismatch repair, are aberrantly expressed during cancer development and progression (6–10). Dysregulated DDR is associated with increased genome instability in HCC cells and has a significant impact on patient prognosis (11).

The use of high-throughput sequencing technology has become increasingly prevalent in recent years, and sequencing data and clinical follow-up information can be downloaded from many cancer databases. In this study, we downloaded the hepatocellular carcinoma dataset from TCGA and GEO databases to explore the prognostic potential of DNA damage response (DDR)-linked genes in hepatocellular carcinoma (HCC) and develop a risk model.

We identified 150 DDR-related genes from the MSigDB database and constructed an 11-gene HCC prognostic signature using univariate Cox regression and random forest analyses. The robustness of the model was validated through internal and external validation. Additionally, we used Gene Set Enrichment Analysis (GSEA) to identify potential pathways associated with the risk model in HCC and analyzed the correlation between clinical traits and the risk score. Finally, we identified and validated two molecular subtypes of HCC using DDR gene expression. Our findings provide novel insights into the molecular mechanisms of HCC and establish an independent DDR gene-based prognostic signature. The use of this signature could aid in personalized therapy and improve clinical decision-making for HCC patients. With the increasing availability of sequencing data, this study provides a useful example of how these data can be utilized to better understand the underlying biology of cancer and improve patient outcomes.

2 Methods

2.1 Data collection

We obtained clinical data and gene expression information for HCC samples from the ICGC-LIRI (<https://dcc.icgc.org/>) and TCGA-LIHC (<https://portal.gdc.cancer.gov/>) datasets. Genes linked to DNA damage response (DDR) were collected from MSigDB, V7.1 (<https://www.gseamsigdb.org/gsea/msigdb>), and only those genes present in both datasets were retained for further analysis.

2.2 Risk signature construction

We utilized univariate and multivariate Cox regression analyses to identify DDR-linked genes in the

LIRI-JP and LIHC datasets. To calculate the risk score for each patient, we used the equation: $(\text{Exp}_i \times \beta_i)$, where Exp_i represents the expression level of prognostic genes and β_i represents the coefficient of Cox regression for each prognostic gene. The median score was used to classify patients into high and low-risk groups, and survival differences were calculated using the “survival” and “survminer” packages. To determine the accuracy of the risk model for 1-, 3-, and 5-year survival, we utilized the “SurvivalROC” package (<https://cran.rproject.org/web/packages/survivalROC/index.html>). We also employed univariate and multivariate Cox regression analyses to determine the prognostic independence of clinical features and the risk score. Potential pathways linked to low and high-risk groups were identified by GSEA, using *c2.cgp.v7.1.symbols.gmt* as the reference gene set.

2.3 Nomogram and DCA curve construction

We constructed a nomogram utilizing independent prognostic factors, and analyzed the benefit of the prognostic factor using decision curve analysis. The discriminative ability of the nomogram was assessed using a calibration plot with the bootstrap approach and 1,000 replications (12). Furthermore, we evaluated the benefit of the prognostic factor using decision curve analysis (13).

2.4 Consensus clustering

We utilized the “ConsensusClusterPlus” R package (with 50 iterations and 80% resampling samples) to group patients into distinct clusters based on DDR-related genes, with the aim of determining molecular subtypes of HCC (14). Principal components analysis (PCA) was employed to distinguish between various LIHC subgroups, and all analyses were validated using the LIRI-JP dataset.

2.5 Immune infiltration analysis

We evaluated the enrichment score of immune cells and infiltration levels of 28 immune cells for each sample in both high- and low-risk groups using the ssGSEA algorithm, which was implemented using the “GSVA” R package (15–17). Furthermore, we analyzed the expression of immune checkpoint genes in both groups.

2.6 Cell culture and transfection

Human HCC cell lines Hep G2 and MHCC-97H were purchased from (National collection of authenticated cell culture,

Shanghai, CN), and incubated at 37 °C with 5% CO₂ in a humidity saturated environment. Cells were cultured in DMEM (Hyclone, LA, USA) and supplied with 10% fetal bovine serum (BI, Israel), anti biotics (0.1 U/l penicillin and 100 g/l streptomycin). DGUOK siRNA were obtained from RiboBio Co., Ltd. (Guangzhou, China). The siRNA was dissolved in DEPC-treated water. Lipofectamine 2000 reagent (Invitrogen, CA, USA) were used for transfection according to the manufacturer's protocol. The solutions were mixed together and incubated at room temperature for 30 minutes. 30 nM siRNA was added into each well and incubated at 37 °C

2.7 Hoechst staining

Cell apoptosis was observed by the morphological changes of the cell nucleus (chromatin agglutination or DNA fragmentation). Cells were treated with si-NC or si-DGUOK, and washed with

PBS twice, Hoechst 33258 (1 µg/ml) was added for 20 min at room temperature avoiding light. Images were gathered by fluorescence microscope (Nikon, Japan)

2.8 Cell viability assays

Cells were seeded in 96-well plates at 10,000 cells per well, and cultured for 24h. They were treated with si-NC or si-DGUOK. Then CCK-8 were added to each plate, absorbance was measured at 450 nm using a FLUOstar Omega microplate reader (BMG Labtech). Cell viability of samples was calculated according to the manufacturer's instructions

2.9 Statistical analyses

The statistical analysis was performed using R (<https://www.r-project.org/>). Kaplan-Meier (KM) method was employed to analyze the survival data and a p-value less than 0.05 was considered statistically significant.

3 Results

3.1 Identification of survival-related DDR risk model

150 DDR-related genes data were obtained from the TCGA dataset, which consisted of 343 HCC samples. Using univariate Cox regression analysis, 37 prognostic genes that affect the survival of patients with HCC were identified. For developing a risk model, stepwise multivariate Cox regression analysis was conducted and 11 genes (AAAS, CANT1, CLP1, DGUOK, GTF2B, GTF2H1, NCBP2, POLA1, POLE4, POLR2D, and POLR2E) were selected. The risk score for each patient was calculated using the following method and computation: AAAS * -0.022 + CANT1 * 0.016 + CLP1 * -0.098 + DGUOK * -0.016 + GTF2B * 0.018 + GTF2H1 * 0.034 + NCBP2 *

0.042 + POLA1 * 0.089 + POLE4 * 0.015 + POLR2D * 0.047 + POLR2E * 0.007.

The patients were then divided into low and high-risk groups based on the median risk score. As shown in Figure 1A, patients in the low-risk group had a longer expected survival rate compared to those in the high-risk group. Furthermore, KM analysis confirmed better prognosis in the low-risk group than in the high-risk group (p < 0.001) (Figure 1B). The predictive performance of the risk model was assessed using ROC analysis, and the area under curve (AUC) values for 1- and 3-year survival were 0.76 and 0.66, respectively (Figure 1C), indicating good accuracy.

3.2 External validation of the DDR-gene prognostic signature

To evaluate the reliability and robustness of the 11-gene signature, we obtained a dataset of 231 HCC samples from ICGC (<https://dcc.icgc.org/>). Risk scores were computed for each patient, and the cohort was divided into high- and low-risk groups. Consistent with the previous findings, the majority of surviving cases were classified into the low-risk group, while a smaller proportion of surviving patients were classified into the high-risk group with higher mortality rate (Figure 2A). KM analysis confirmed better survival outcomes for individuals in the low-risk group compared to their high-risk counterparts (Figure 2B). The AUC values for 1- and 3-year survival were 0.77 and 0.73, respectively, indicating a good prognostic performance of the risk model in HCC (Figure 2C).

3.3 The risk model is an independent prognostic predictor of HCC

Univariate and multivariate Cox regression analyses showed that the risk model and tumor stage were independent risk factors for HCC, as reported in Figures 3A, B. Furthermore, ROC analysis demonstrated that the risk model performed better than tumor stage in predicting 1-year prognosis, with AUC values of 0.746 and 0.700, respectively (Figure 4A). Subsequently, we constructed a nomogram that integrated the risk model and tumor stage to predict overall survival (OS) at 1-, 3-, and 5-year timepoints (Figure 5A). The nomogram exhibited good prognostic performance, as indicated by AUC values at 1-, 3-, and 5-year timepoints (Figures 4B, C). The stability of the nomogram was further validated by calibration curve plots (Figures 5B–D). Overall, the DDR gene-based risk score and tumor stage-based nomogram can robustly predict the prognosis of HCC patients and thus, can be useful in clinical decision-making.

3.4 Gene set enrichment analysis

To identify enriched pathways in HCC, GSEA was conducted for both high- and low-risk groups.

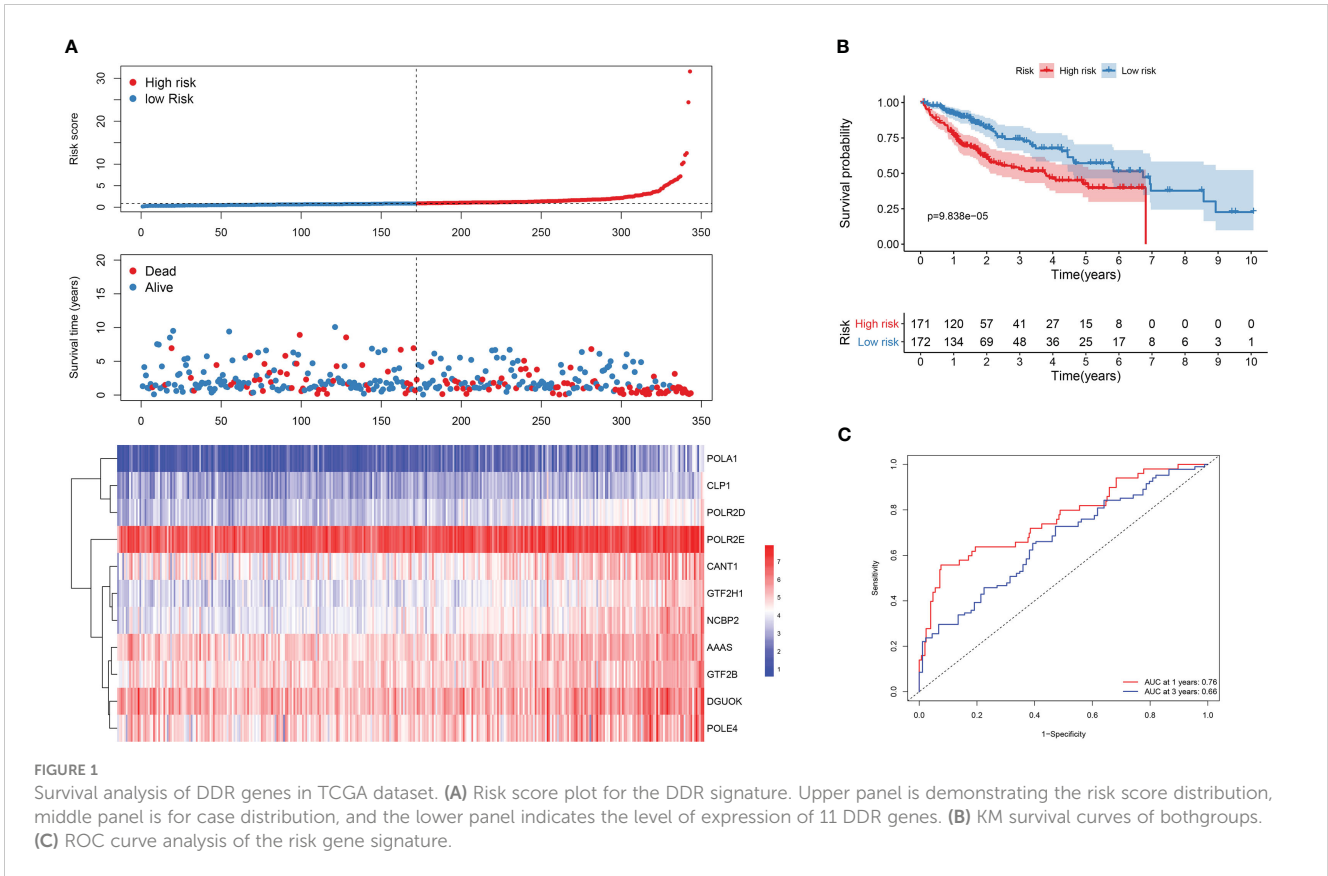


FIGURE 1 Survival analysis of DDR genes in TCGA dataset. **(A)** Risk score plot for the DDR signature. Upper panel is demonstrating the risk score distribution, middle panel is for case distribution, and the lower panel indicates the level of expression of 11 DDR genes. **(B)** KM survival curves of both groups. **(C)** ROC curve analysis of the risk gene signature.

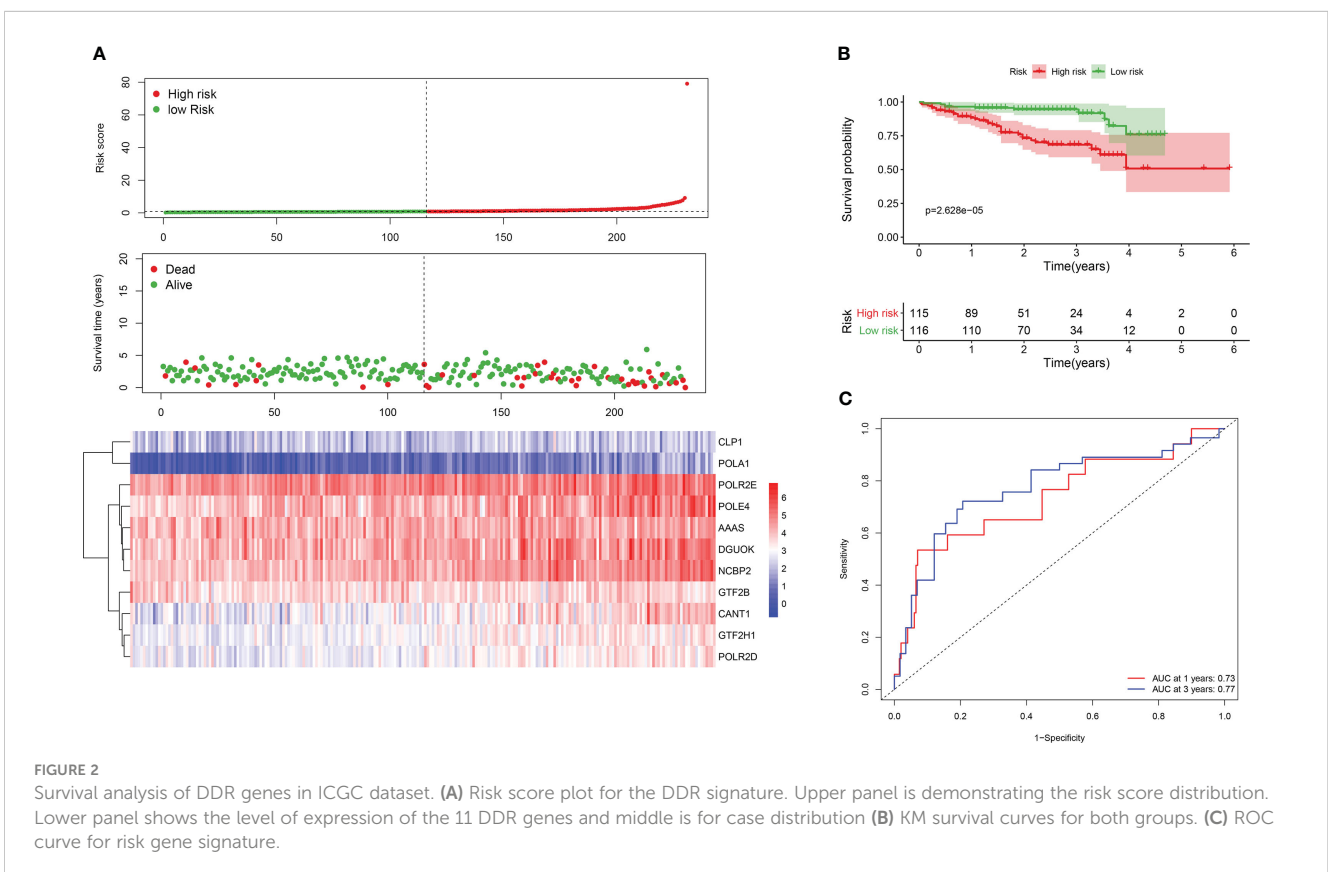


FIGURE 2 Survival analysis of DDR genes in ICGC dataset. **(A)** Risk score plot for the DDR signature. Upper panel is demonstrating the risk score distribution. Lower panel shows the level of expression of the 11 DDR genes and middle is for case distribution **(B)** KM survival curves for both groups. **(C)** ROC curve for risk gene signature.

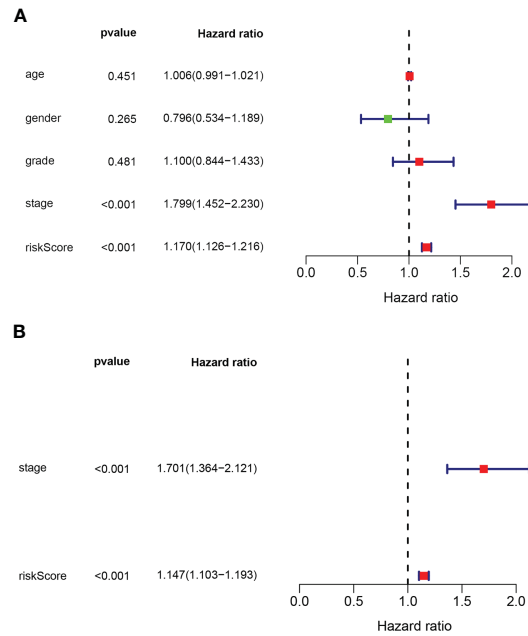


FIGURE 3 (A) Univariate and (B) multivariate cox regression analyses used to get prognostic value of the gene biosignature and clinical traits.

Results showed that pathways associated with apoptosis, cell cycle, and MAPK, mTOR, NOCTH, UBIQUITIN, and WNT signaling were enriched in the high-risk group, while the low-risk group exhibited enrichment of pathways related to metabolism of fatty acid and retinol (Figures 6A, B). These findings suggest that favorable prognosis and low-risk scores are correlated with metabolism-linked pathways, whereas cancer-related pathways coincide with high-risk scores and poor prognosis.

distribution function curve and the consensus heatmap (Figures 7A-C). PCA further differentiated patients into two distinct subgroups (Figure 7D), with subgroup 1 exhibiting better overall survival compared to subgroup 2 (Figure 7E). To validate the robustness of the classification, we also evaluated these subgroups in the ICGC dataset (Figure 8). Further analysis of the correlation and clinical characteristics in both TCGA and ICGC datasets for the two subgroups showed that the group with better survival outcomes included more early-stage cases (Figures 9A, B).

3.5 Identification of molecular subtypes of HCC

Consensus clustering algorithm based on the DDR genes was employed to identify the molecular subtypes of HCC. The optimal cluster was determined to be K=2 using the cumulative

3.6 Correlation of the immune infiltration with HCC subclasses

The ssGSEA algorithm was used to analyze the infiltration of 24 immune cells in both the high- and low-risk groups. The high-risk

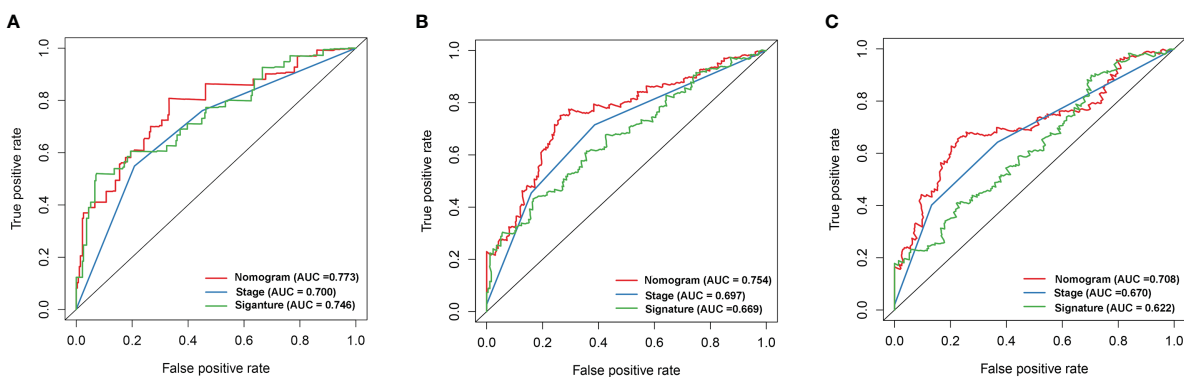


FIGURE 4 The gene signature, ROC curve analysis of the nomogram, and disease stage for 1- (A), 3- (B), and 5-year (C) survival.

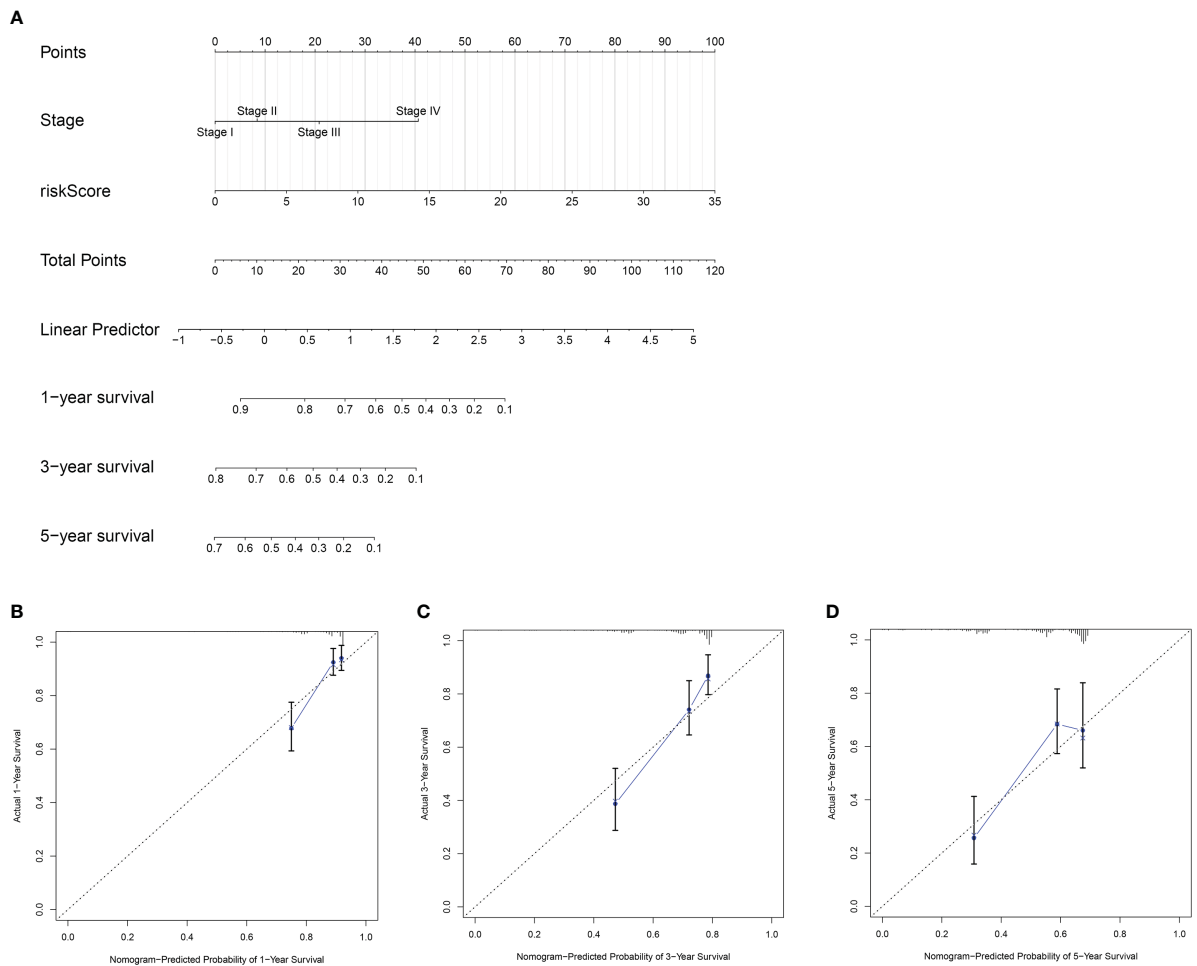


FIGURE 5 Construction and validation of a nomogram(prognostic). **(A)** A nomogram based on the risk model and tumor stage to estimate overall survival of HCC patients. The estimating of 1- **(B)**, 3- **(C)**and 5-year **(D)** survival of HCC patients using calibration curve plot of the nomogram.

group showed high levels of infiltration of activated CD4+ T cells, CD4+ T (central memory) CD8+ T cells (central memory), CD4+ T cells (effector memory), B cells(memory), regulatory T cells, T follicular helper cells, Th17 cells, Th2 cells, activated CD8+ T cells,immature dendritic cells (DCs), and plasmacytoid DCs,

whereas the low-risk group had greaterinfiltration of activated CD8+ T cells and eosinophils (**Figure 10A**). Additionally, the high-risk group exhibited comparatively higher expression of all inhibitory immune receptors compared to the low-risk group (**Figure 10B**). These findings suggest that the anti-tumor

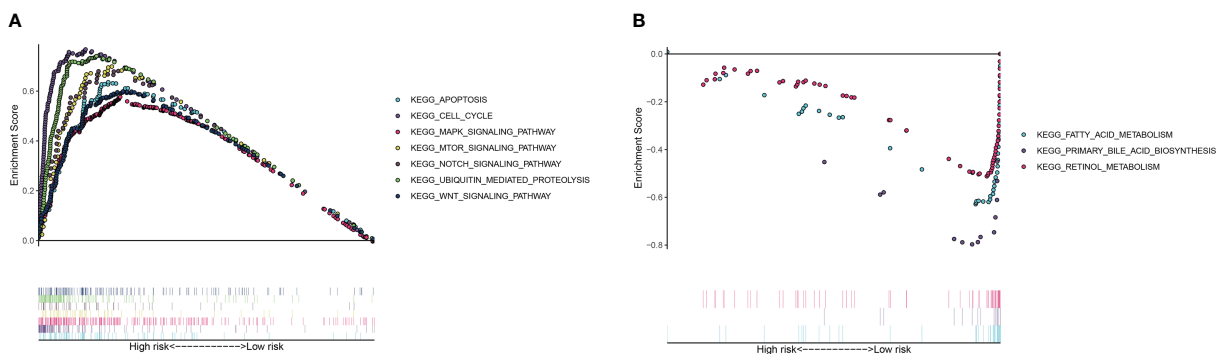
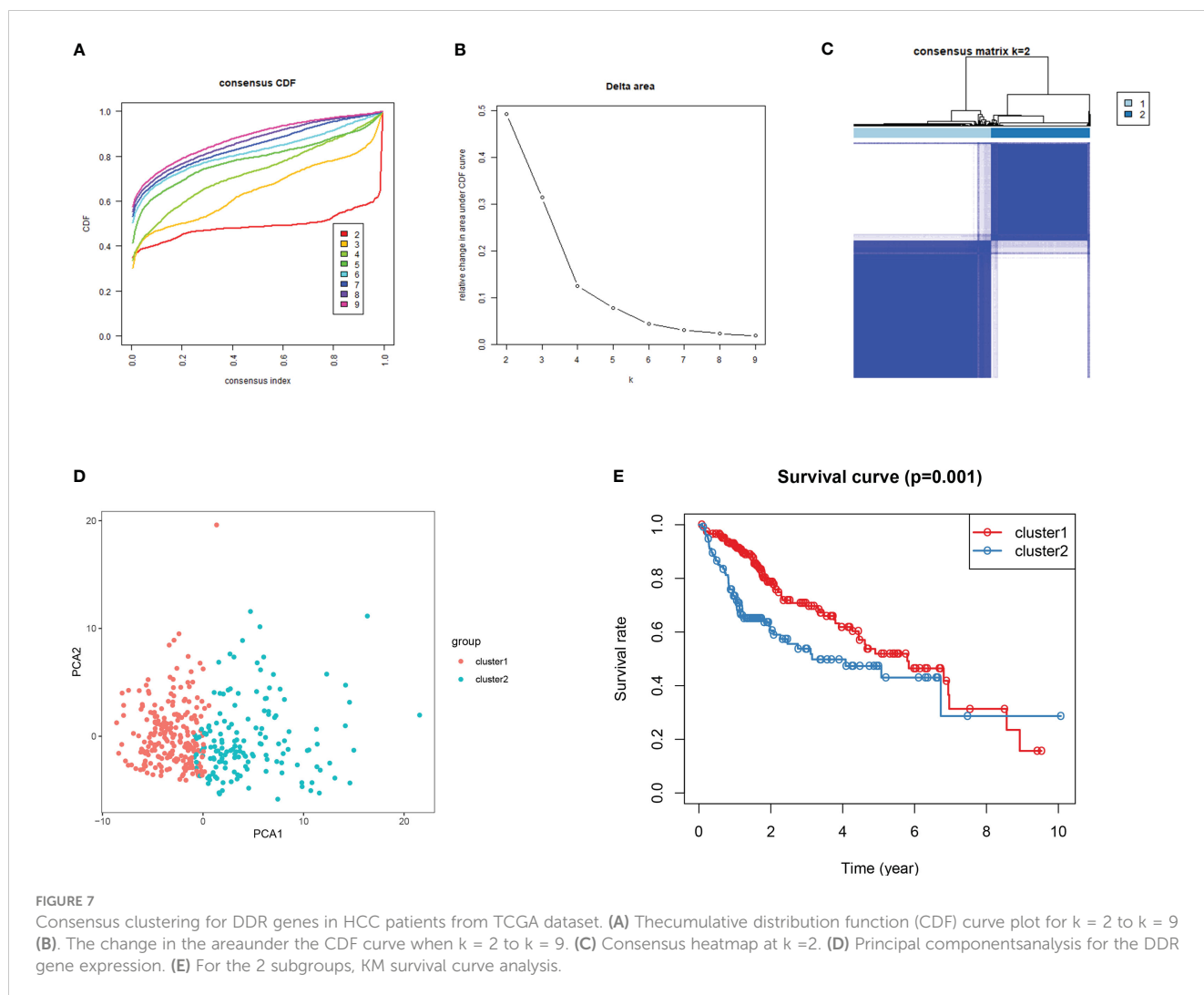


FIGURE 6 Gene set enrichment analysis results showing enriched pathways in **(A)** high-risk and **(B)** low-risk groups.



properties of high T cell infiltration were offset by a strong immunosuppressive tumor microenvironment due to the overexpression of immune checkpoint proteins (18).

3.7 Cell assays

In vitro validation on DGUOK. HepG2 and MHCC-97H cells were treated with CCK8 and were performed to detect the cell viability. Hoechst 33258 fluorescent dye staining was used to show nuclear morphological changes and to assess apoptosis. Data in A are presented as means \pm SD. **, $p < 0.01$. magnification: 200 \times . The results showed the proliferation ability is higher in the si-NC group compared by the si-DGUOK group (Figure 11).

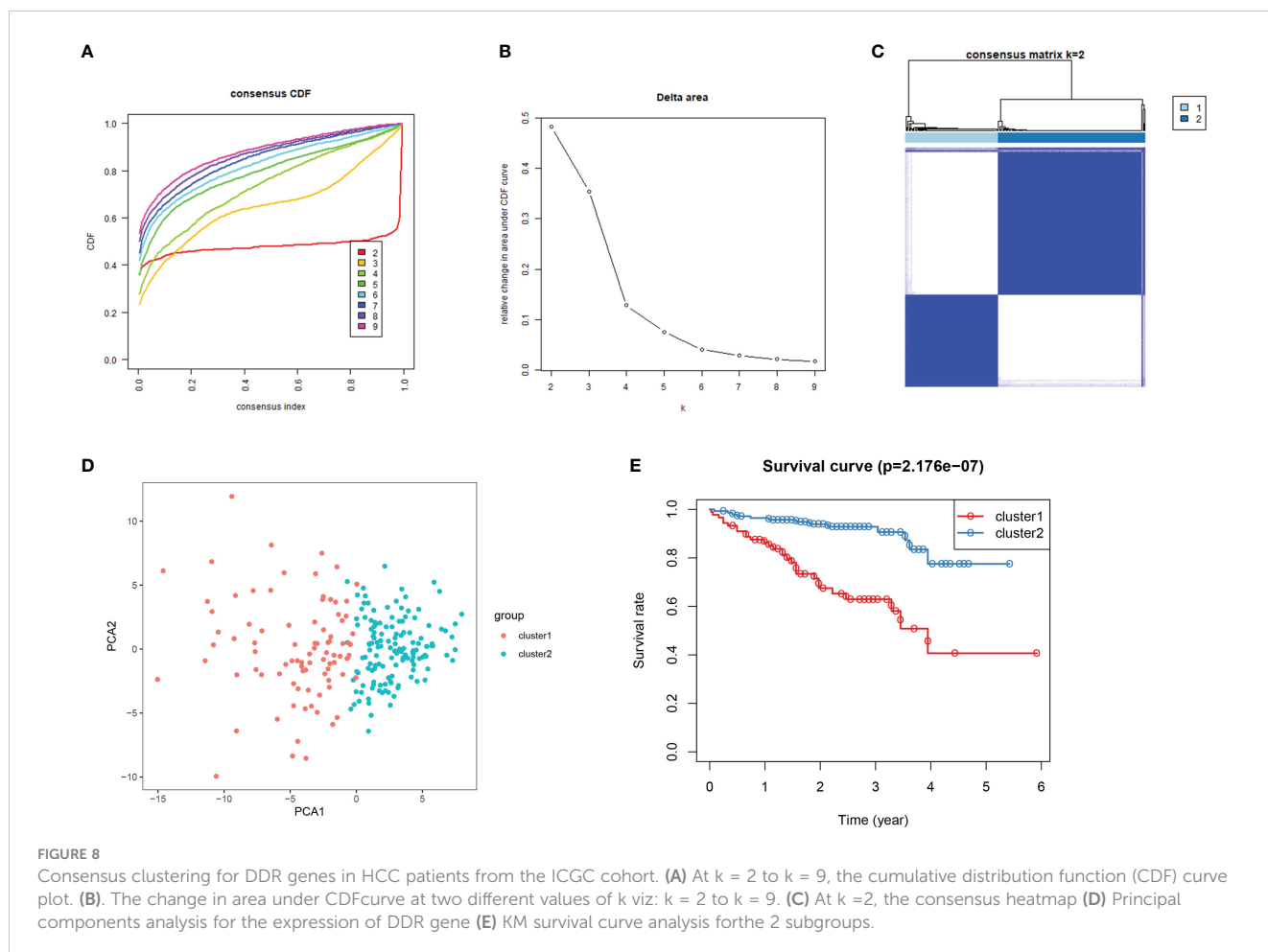
3.8 MiRNA-mRNA Network

Based on the targets database data, we performed a relevant microRNA analysis of DNA repair genes in Siganture using the

Cytoscape software, and conducted DNA repair related miRNA-mRNA Network (Figure 12).

4 Discussion

HCC is a highly heterogeneous cancer with multiple risk factors, including alcohol consumption, hepatitis B/C infection, and obesity (19). The initiation of HCC is associated with DNA damage and chromosomal abnormalities, which triggers a DNA damage response (DDR) in affected cells. DNA lesions can be repaired through various mechanisms, including homologous recombination, mismatch repair, and double-strand break repair. Dysfunctional DDR pathways can result in genomic instability, mutations, and eventually lead to HCC development and progression (20). Many DNA repair proteins, such as sphingolipid signaling, TP53, hOGG1, XRCC1, PARP-1, MRE11-Rad50NBS1 (MRN) complex, and ataxia-telangiectasia mutant (ATM) kinase, are frequently mutated in HCC (21). Furthermore, ionizing radiation (IR)-induced DDR pathways can create an



immunosuppressive tumor microenvironment, thereby weakening the anti-tumor effect of radioimmunotherapy. DDR inhibitors can reverse the immunosuppressive state of HCC and inhibit tumor progression, providing a potential therapeutic strategy (22).

There is a substantial body of evidence suggesting that DNA damage response (DDR) genes play a crucial role in the development of cancer (23). DDR genes are often expressed abnormally in mucosal or tumor tissues and are closely associated with patient prognosis (24, 25). However, the ability of individual genes to serve as prognostic markers is limited (26, 27), and multi-gene signatures may be better suited for predicting the prognosis of hepatocellular carcinoma (HCC). Despite this, no study has yet investigated the prognostic value of DDR genes in HCC. To address this gap, we developed an 11-gene signature based on DDR gene expression data and clinical data obtained from the ICGC and TCGA databases. The risk score generated by the 11-gene signature enabled the classification of patients into low- and high-risk groups, with the latter group exhibiting poorer survival outcomes. The risk model demonstrated good predictive performance in both TCGA and ICGC datasets. Additionally, the risk model was found to be an independent prognostic factor for HCC. A nomogram constructed using the risk score and tumor stage allowed the clear differentiation of two prognostic groups, which may be helpful in guiding preoperative management of HCC patients. The DDR gene

signature identified in this study was found to be linked with several cancer related pathways including cell cycle, WNT signaling, mTOR signaling and apoptosis in the high risk group, which may be indicative of the potential mechanisms underlying HCC progression. On the other hand, the low-risk group was enriched in metabolism-related pathways. Most of the genes in the DDR-based risk signature have been implicated in tumorigenesis. For instance, CANT1 is known to regulate pyrimidine metabolism in melanoma cells and is associated with tumor progression (28). High expression of CANT1 in prostate cancer cells has been associated with better prognosis, while its silencing significantly suppressed cell proliferation and DNA synthesis (29). CLP1, on the other hand, plays an important role in motor neuron function (30). Mitochondrial deoxyguanosine kinase (DGUOK) is an enzyme that controls the rate of deoxy nucleoside salvage pathway in the mitochondria. Overexpression of DGUOK has been associated with worse prognosis in lung cancer, and its depletion suppressed lung adenocarcinoma growth, CSC self-renewal and metastasis (31). GTF2B has been identified as a prognostic marker for colorectal cancer and neuroblastoma, while GTF2H1 is a p62 subunit of complex transcription factor IIIH (TFIIH) that regulates nucleotide excision repair and transcription (32, 33). Certain polymorphisms/haplotypes of GTF2H1 have been associated with increased susceptibility to

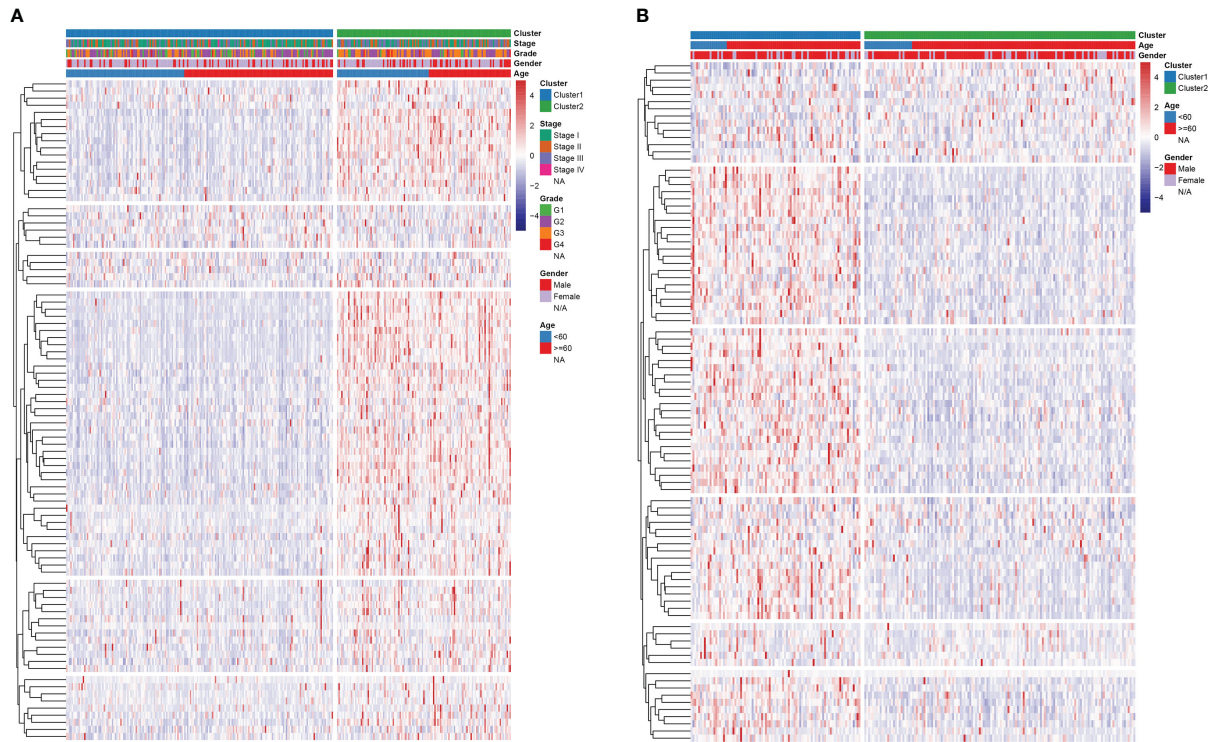


FIGURE 9 Heatmap analysis of the relationship between subgroup and clinical traits in (A) TCGA and (B) ICGC cohorts.

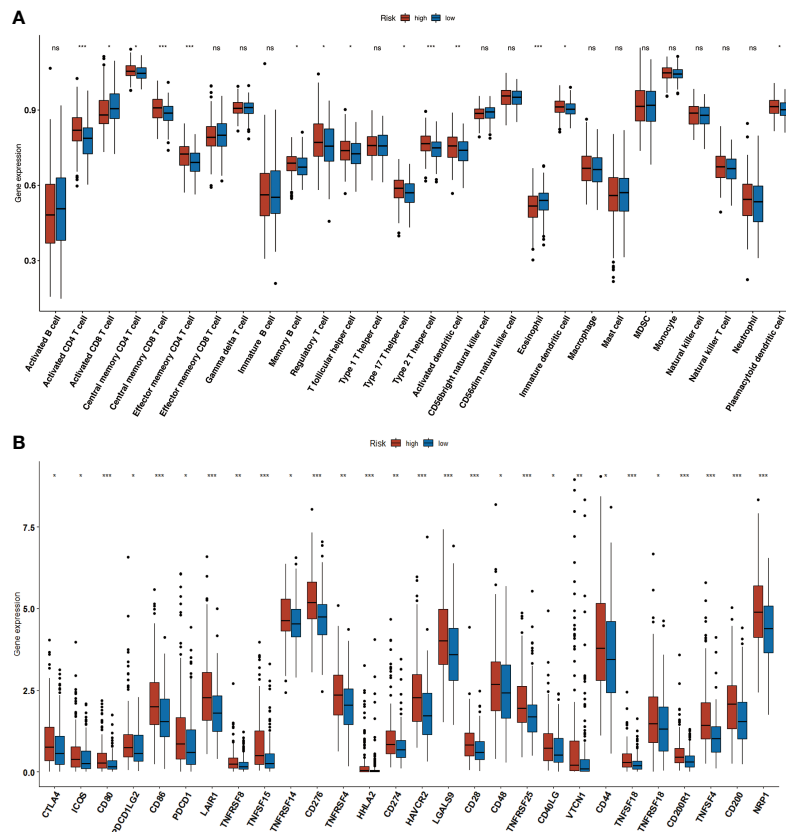


FIGURE 10 Immune infiltration. (A) The increased number of tumor infiltrating immune cells in high- and low-risk groups. (B) Boxplot showing the immune-checkpoint genes expression in low- and high-risk groups. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

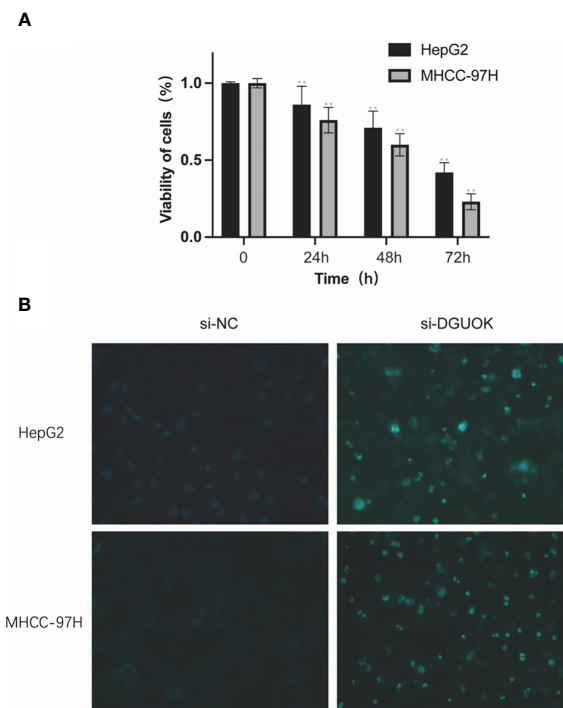


FIGURE 11
In vitro validation on DGUOK. HepG2 and MHCC-97H cells were treated with siDGUOK for indicated time. **(A)**: CCK8 was performed to detect the cell viability. **(B)**: Hoechst 33258 fluorescent dye staining was used to show nuclear morphological changes and to assess apoptosis. Data in **(A)** are presented as means \pm SD. ** $p < 0.01$. magnification: 200 \times .

lung cancer (34). Additionally, the budding yeast orthologs of POLE4 have been shown to enhance Pole processivity *in vitro*, but have the opposite effect *in vivo*, leading to accelerated tumorigenesis (35).

Moreover, the POLR2E rs3787016 polymorphism may enhance the risk of developing the prostate cancer, liver cancer esophageal cancer, papillary thyroid carcinoma, and breast cancer (36–38). However, the role of AAAS, NCBP2, POLA1 and POLR2D in HCC

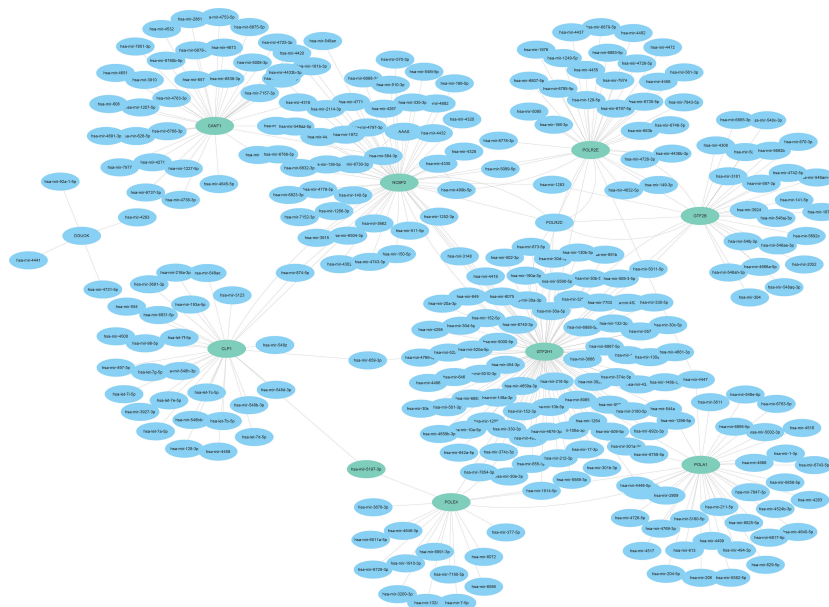


FIGURE 12
 DNA Repair Gene miRNA-mRNA Network.

is unknown, and will have to be experimentally verified. Immunotherapy has achieved encouraging results in various malignancies (39), including HCC. For instance, the “T+A” scheme is increasingly becoming the first-line option for advanced HCC (40). Despite achieving good outcomes in multiple cancers, a significant percentage of the patients do not benefit from immunotherapy. Hence it is necessary to recognize the biomarkers that can reveal the outcomes of immunotherapy, and screen for patients that can respond to immunotherapeutic regimens. Galon et al. (41) had proposed the concept of “cold” and “hot” tumors to evaluate their sensitivity to immunotherapies. In this study, we detected increased infiltration of immunosuppressive cells and overexpression of receptors responsible for immune checkpoint in the high-risk group, which indicates that the high-risk group patients are likely unresponsive to immunotherapy.

In summary, we identified biomarkers of HCC based on computational biology in oncology methods (42, 43), and constructed prognostic models using machine learning methods (44–46). We have established an 11-DDR gene signature that can accurately forecast the prognosis of hepatocellular carcinoma (HCC). The utilization of this prognostic signature not only advances our comprehension of the underlying molecular mechanisms that contribute to HCC progression but also provides a practical guide for clinical decision-making.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material. Further inquiries can be directed to the corresponding author.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* (2018) 68(6):394–424. doi: 10.3322/caac.21492
2. Davis GL, Dempster J, Meler JD, Orr DW, Walberg MW, Brown B, et al. Hepatocellular carcinoma: management of an increasingly common problem. *Proc (Bayl Univ Med Cent)* (2008) 21(3):266–80. doi: 10.1080/08998280.2008.11928410
3. Marengo A, Rosso C, Bugianesi E. Liver cancer: connections with obesity, fatty liver, and cirrhosis. *Annu Rev Med* (2016) 67:103–17. doi: 10.1146/annurev-med-090514-013832
4. McGlynn KA, Petrick JL, London WT. Global epidemiology of hepatocellular carcinoma: an emphasis on demographic and regional variability. *Clin Liver Dis* (2015) 19(2):223–38. doi: 10.1016/j.cld.2015.01.001
5. Lans H, Hoeijmakers JHJ, Vermeulen W, Marteijn JA. The DNA damage response to transcription stress. *Nat Rev Mol Cell Biol* (2019) 20(12):766–84. doi: 10.1038/s41580-019-0169-4
6. Nickoloff JA, Jones D, Lee SH, Williamson EA, Hromas R. Drugging the cancers addicted to DNA repair. *J Natl Cancer Inst* (2017) 109(11). doi: 10.1093/jnci/djx059
7. Laurini E, Marson D, Fermeglia A, Aulic S, Fermeglia M, Prisl S. Role of Rad51 and DNA repair in cancer: a molecular perspective. *Pharmacol Ther* (2020) 208:107492. doi: 10.1016/j.pharmthera.2020.107492
8. Barretti M, Le DT. DNA Mismatch repair in cancer. *Pharmacol Ther* (2018) 189:45–62. doi: 10.1016/j.pharmthera.2018.04.004
9. Mateo J, Boyens G, Barbieri CE, Bryant HE, Castro E, Nelson PS, et al. DNA Repair in prostate cancer: biology and clinical implications. *Eur Urol* (2017) 71(3):417–25. doi: 10.1016/j.eururo.2016.08.037
10. Wang X, Tan C, Ye M, Wang X, Weng W, Zhang M, et al. Development and validation of a DNA repair gene signature for prognosis prediction in colon 334 cancer. *J Cancer* (2020) 11(20):5918–28. doi: 10.7150/jca.46328
11. Lin J, Shi J, Guo H, Yang X, Jiang Y, Long J, et al. Alterations in DNA damage repair genes in primary liver cancer. *Clin Cancer Res* (2019) 337 25(15):4701–11. doi: 10.1158/1078-0432.CCR-19-0127
12. Iasonos A, Schrag D, GV R, Panageas KS: how to build and interpret a nomogram for cancer prognosis. *J Clin Oncol* (2008) 26(8):1364–70. doi: 10.1200/JCO.2007.12.9791
13. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* (2006) 26(6):565–74. doi: 10.1177/0272989X06295361
14. Wilkerson MD, Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. *Bioinformatics* (2010) 26(12):1572–3. doi: 10.1093/bioinformatics/btq170
15. Hanzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf* (2013) 14:7. doi: 10.1186/1471-2105-14-7
16. Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, et al. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep* (2017) 18(1):248–62. doi: 10.1016/j.celrep.2016.12.019
17. Jia Q, Wu W, Wang Y, Alexander PB, Sun C, Gong Z, et al. Local mutational diversity drives intratumoral immune heterogeneity in non-small cell lung cancer. *Nat Commun* (2018) 9(1):5361. doi: 10.1038/s41467-018-07767-w
18. Matsushita H, Sato Y, Karasaki T, Nakagawa T, Kume H, Ogawa S, et al. Neoantigen load, antigen presentation machinery, and immune signatures determine

Ethics statement

The studies involving human participants were reviewed and approved by Intensive Care Medicine, Liaoyou Baoshihua Hospital. The patients/participants provided their written informed consent to participate in this study.

Author contributions

YB contributed to data analysis, methodology, figures construction, and article writing. YW contributed to investigation and validation. JH, YM, HL and ML contributed to methodology and validation. All authors contributed to the article and approved the submitted version.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- prognosis in clear cell renal cell carcinoma. *Cancer Immunol Res* (2016) 4(5):463–71. doi: 10.1158/2326-6066.CIR-15-0225
19. Villanueva A. Hepatocellular carcinoma. *N Engl J Med* (2019) 380(15):1450–62. doi: 10.1056/NEJMra1713263
20. O'Connor MJ. Targeting the DNA damage response in cancer. *Mol Cell* (2015) 60(4):547–357 560. doi: 10.1016/j.molcel.2015.10.040
21. Yang SF, Chang CW, Wei RJ, Shiue YL, Wang SN, Yeh YT. Involvement of DNA damage response pathways in hepatocellular carcinoma. *BioMed Res Int* (2014) 2014:153867. doi: 10.1155/2014/153867
22. Sheng H, Huang Y, Xiao Y, Zhu Z, Shen M, Zhou P, et al. ATR inhibitor AZD6738 enhances the antitumor activity of radiotherapy and immune checkpoint inhibitors by potentiating the tumor immune microenvironment in hepatocellular carcinoma. *J Immunother Cancer* (2020) 8(1). doi: 10.1136/jitc-2019-000340
23. Ali R, Rakha EA, Madhusudan S, Bryant HE. DNA Damage repair in breast cancer and its therapeutic implications. *Pathology* (2017) 49(2):156–65. doi: 10.1016/j.pathol.2016.11.002
24. Jinjia C, Xiaoyu W, Hui S, Wenhua L, Zhe Z, Xiaodong Z, et al. The use of DNA repair genes as prognostic indicators of gastric cancer. *J Cancer* (2019) 10(20):4866–75. doi: 10.7150/jca.31062
25. Kun S, Duan Q, Liu G, Lu JM. Prognostic value of DNA repair genes based on stratification of glioblastomas. *Oncotarget* (2017) 8(35):58222–30. doi: 10.18632/oncotarget.17452
26. Deng CX, Wang RH. Roles of BRCA1 in DNA damage repair: a link between development and cancer. *Hum Mol Genet* (2003) 12 Spec No 1:R113–123. doi: 10.1093/hmg/ddg082
27. Fang M, Xia F, Mahalingam M, Virbasius CM, Wajapeyee N, Green MR. MEN1 is a melanoma tumor suppressor that preserves genomic integrity by stimulating transcription of genes that promote homologous recombination-directed DNA repair. *Mol Cell Biol* (2013) 33(13):2635–375 2647. doi: 10.1128/MCB.00167-13
28. Edwards L, Gupta R, Filipp FV. Hypermutation of DPYD deregulates pyrimidine metabolism and promotes malignant progression. *Mol Cancer Res* (2016) 14(2):196–206. doi: 10.1158/1541-7786.MCR-15-0403
29. Gerhardt J, Steinbrech C, Buchi O, Behnke S, Bohnert A, Fritzsche F, et al. The androgen-regulated calcium-activated nucleotidase 1 (CANT1) is commonly overexpressed in prostate cancer and is tumor-biologically relevant *in vitro*. *Am J Pathol* (2011) 178(4):1847–60. doi: 10.1016/j.ajpath.2010.12.046
30. Hanada T, Weitzer S, Mair B, Bernreuther C, Wainger BJ, Ichida J, et al. Komnenovic V et al: CLP1 links tRNA metabolism to progressive motor-neuron loss. *Nature* (2013) 495(7442):474–80. doi: 10.1038/nature11923
31. Lin S, Huang C, Sun J, Bolt O, Wang X, Martine E, et al. The mitochondrial deoxyguanosine kinase is required for cancer cell stemness in lung adenocarcinoma. *EMBO Mol Med* (2019) 11(12):e10849. doi: 10.15252/emmm.201910849
32. Di Pietro C, Ragusa M, Barbagallo D, Duro LR, Guglielmino MR, Majorana A, et al. Involvement of GTA protein NC2beta in neuroblastoma pathogenesis suggests that it physiologically participates in the regulation of cell proliferation. *Mol Cancer* (2008) 7:52. doi: 10.1186/1476-4598-7-52
33. Ribeiro-Silva C, Aydin OZ, Mesquita-Ribeiro R, Slyskova J, Helfricht A, Martein JA, et al. DNA Damage sensitivity of SWI/SNF-deficient cells depends on TFIIH subunit p62/GTF2H1. *Nat Commun* (2018) 9(1):4067. doi: 10.1038/s41467-018-06402-y
34. Wu W, Liu H, Lei R, Chen D, Zhang S, Lv J, et al. Genetic variants in GTF2H1 and risk of lung cancer: a case-control analysis in a Chinese population. *Lung Cancer* (2009) 63(2):180–6. doi: 10.1016/j.lungcan.2008.05.014
35. Bellelli R, Borel V, Logan C, Svendsen J, Cox DE, Nye E, et al. Polepsilon instability drives replication stress, abnormal development, and tumorigenesis. *Mol Cell* (2018) 70(4):707–721 e707.
36. Chen B, Li J, Yi C, Jiao Y, Gu X, Feng X. Long non-coding RNA POLR2E rs3787016 is associated with the risk of papillary thyroid carcinoma in Chinese population. *Pathol Res Pract* (2018) 214(7):1040–4.
37. Chen B, Jiao Y, Yaolong F, Li T, Liu Y, Wang M, et al. The POLR2E rs3787016 polymorphism is strongly associated with the risk of female breast and cervical cancer. *Pathol Res Pract* (2019) 215(5):1061–5.
38. Chen B, Wang S, Ma G, Han J, Zhang J, Gu X, et al. The association of POLR2E rs3787016 polymorphism and cancer risk: a Chinese case-control study and meta-analysis. *Biosci Rep* (2018) 38(6).
39. Yang Y. Cancer immunotherapy: harnessing the immune system to battle cancer. *J Clin Invest* (2015) 125(9):3335–7.
40. Finn RS, Qin S, Ikeda M, PR G, Ducreux M, TY K, et al. Atezolizumab plus bevacizumab in unresectable hepatocellular carcinoma. *N Engl J Med* (2020) 382(20):1894–905.
41. Galon J, Bruni D. Approaches to treat immune hot, altered and cold tumours with combination immunotherapies. *Nat Rev Drug Discovery* (2019) 18(3):197–218. doi: 10.1038/s41573-018-0007-y
42. Wang Y, Yan K, Lin J, Li J, Bi J. Macrophage M2 Co-expression factors correlate with the immune microenvironment and predict outcome of renal clear cell carcinoma. *[[[Front Genet* (2021) 12:615655. doi: 10.3389/fgene.2021.615655
43. Wang Y, Yan K, Wang J, Lin J, Bi J. M2 macrophage Co-expression factors correlate with immune phenotype and predict prognosis of bladder cancer. *Front Oncol* (2021) 11:609334(undefined). doi: 10.3389/fonc.2021.609334
44. Wang Y, Yan K, Wang L, Bi J. Genome instability-related long non-coding RNA in clear renal cell carcinoma determined using computational biology. *[[[BMC Cancer* (2021) 21:727. doi: 10.1186/s12885-021-08356-9
45. Wang L, Shen J, Wang Y, Bi J. Identification of fatty acid metabolism-based molecular subtypes and prognostic signature to predict immune landscape and guide clinical drug treatment in renal clear cell carcinoma. *[[[Int Immunopharmacol* (2023) 116:109735. doi: 10.1016/j.intimp.2023.109735
46. Hu J, Wang L, Li L, Wang Y, Bi J. A novel focal adhesion-related risk model predicts prognosis of bladder cancer – a bioinformatic study based on TCGA and GEO database. *[[[BMC Cancer* (2022) 22:1158. doi: 10.1186/s12885-022-10264-5