Original Article NABP2 as an oncogenic biomarker promotes hepatocellular carcinoma progression and metastasis

Hong Peng^{4*}, Shenglan Cai^{3,5,6*}, Ruochan Chen^{5,6*}, Linxia Tan^{5,6}, Shanshan Lu^{6,7}, Xingjun Lu², Yuanxin Hang², Yiya Zhang^{1,5,8}, Xiaozhen Peng^{1,2,9}

¹Department of Dermatology, Xiangya Hospital, Central South University, Changsha, Hunan, China; ²School of Public Health & Laboratory Medicine, Hunan University of Medicine, Huaihua, Hunan, China; ³Infection Control Department, Hunan Provincial Maternal and Child Health Care Hospital, Changsha, Hunan, China; ⁴Medical School, Huanghe Science & Technology College, Zhengzhou, Henan, China; ⁵Hunan Key Laboratory of Viral Hepatitis, Xiangya Hospital, Central South University, Changsha, Hunan, China; ⁶National Clinical Research Center for Geriatric Disorders, Xiangya Hospital, Central South University, Changsha, Hunan, China; ⁷The Higher Educational Key Laboratory for Cancer Proteomics and Translational Medicine of Hunan Province, Xiangya Hospital, Central South University, Changsha, Hunan, China; ⁸Hunan Key Laboratory of Aging Biology, Xiangya Hospital, Central South University, Changsha, Hunan, China; ⁹Hunan Provincial Key Laboratory for Synthetic Biology of Traditional Chinese Medicine, Hunan University of Medicine, Huaihua, Hunan, China. *Equal contributors.

Received December 5, 2022; Accepted May 13, 2023; Epub June 15, 2023; Published June 30, 2023

Abstract: Objectives: To evaluate the role and biological function of nucleic acid binding protein 2 (NABP2) in hepatocellular carcinoma (HCC). Methods: Our study was based on comprehensive bioinformatics methods and functional analysis experiments using HCC cells to reveal the expression of NABP2, the prognostic role of NABP2, the relationship between NABP2 and the infiltration of immune cells and the expression of immune-related cytokines, potential effective drugs against HCC, and the biological function of NABP2 in HCC. Results: Our results indicated that NABP2 expression was markedly elevated in HCC, which suggested a worse prognosis and shorter survival time in HCC patients. Moreover, NABP2 was an independent prognostic factor and was associated with cancer-related signal pathways in HCC. Further functional analysis showed that knockdown of NABP2 dramatically inhibited proliferation and migration, and promoted apoptosis of HCC cells. Subsequently, we identified NABP2-related genes and NABP2related clusters. Next, we constructed a NABP2-related risk signature based on differentially expressed genes that were responsible for NABP2-related clusters. We found that the risk signature was an independent prognostic factor for patients with HCC that was associated with dysregulated immune infiltration. Finally, drug sensitivity analysis revealed eight potentially effective drugs for beneficial treatment options for HCC patients with high-risk scores. Conclusions: These findings indicated that NABP2 is a prognostic biomarker and therapeutic target for HCC, and a NABP2-related risk signature could guide clinicians to judge the prognosis and suggest drug treatments for HCC patients.

Keywords: NABP2, biomarker, risk signature, therapeutic target, immune infiltration, HCC

Introduction

Hepatocellular carcinoma is the most common type of liver cancer and the world's second leading etiology for cancer-associated death [1, 2]. HCC patients have a poor prognosis and high morbidity rate, with a median survival time of only a few months and a 3-year survival rate of 12.7% [3]. For early-stage HCC, surgical resection followed by chemotherapy may be the most effective treatment. However, most patients are diagnosed at an advanced stage at the time of their first presentation and, with currently no effective treatment, their prognosis is poor. With progress in biotechnology and new insights into the mechanism of cancer initiation, molecular-targeted therapy is being developed for HCC treatment [4]. Therefore, the search for novel therapeutic targets and prognostic indicators for patients with HCC is imperative.

Single-strand DNA (ssDNA)-binding proteins (SSBs) are widespread in cells and are essen-

tial for cellular metabolism. SSPs have been implicated in processes such as DNA replication, recombination, DNA damage detection, DNA repair, and RNA transcription [5]. Appropriate DNA manipulation is essential for maintaining genomic integrity and tumor suppression by SSBs [6, 7]. Recently, the novel human single-stranded DNA binding protein NABP2 has been shown to play a critical role in maintaining genomic stability [8]. Notably, NABP2 has been importantly implicated in regulating DNA damage checkpoints and the cell cycle by binding and protecting p53 and p21 proteins from ubiquitin-mediated degradation [9, 10]. Nevertheless, a more comprehensive understanding of NABP2's expression and role in HCC is needed.

In our study, the expression, biological function, and prognostic role of NABP2 in HCC were intensively studied using comprehensive bioinformatics techniques and functional analysis experiments. Moreover, we developed a NABP2-related risk signature based on NABP2 cluster-related genes and established a sensitive and specific nomogram for HCC prognosis predictions. Drug sensitivity analysis was utilized to screen out potentially effective drugs targeting NABP2 for treating HCC patients with high risk. In summary, NABP2 is a novel candidate for a prognostic biomarker as well as a therapeutic target for HCC.

Materials and methods

NABP2 expression and survival analysis

Data on the expression of NABP2 in HCC were obtained from the International Cancer Genome Consortium (ICGC) datasets and the Cancer Genome Atlas (TCGA) using R software. To confirm transcription levels of NABP2 genes in HCC, samples from the Genotype-Tissue Expression (GTEx) project and TCGA project were also evaluated in the Gene Expression Profiling Interactive Analysis (GEPIA) database. Moreover, the Gene Expression Omnibus (GEO) datasets with both HCC tissues and normal liver tissues detected by Affymetrix Human Genome Array were selected as validation sets for NABP2 expression (GSE45436, GSE62232).

GTEx samples were obtained from the UCSC Xena project [phenotype (n = 9,783) and TOIL RSEM from fpm (n = 7,862)] (http://xena.ucsc.

edu/). The official symbol of the gene was derived from the ensembl number of the gene in RNAseq data by our own Perl script, which was applied to obtain NABP2 expression in many normal organs of different genders. Based on the GitHub website R package, gganatogram and ggpubr visualization were conducted.

Immunohistochemistry (IHC) from the Human Protein Atlas (HPA) database was used to analyze NABP2 protein expression levels in HCC. The protein levels of NABP2 from the Clinical Proteomic Tumor Analysis Consortium (CPTAC) proteomics database were applied to the expression of NABP2 protein and prognosis in HCC via the "Limma" package and "survival" package of R software.

From 2015 to 2019, a group of 15 formalinfixed, paraffin-embedded HCCs with adjacent liver tissue samples was provided by the Xiangya Hospital of Central South University. Study approval was received from the Ethics Committee of Xiangya Hospital, Central South University.

Formalin-fixed and paraffin-embedded tissue sections were subjected to IHC and immunoreactive scoring for NABP2 (Affinity Biosciences, DF13202, 1:100), as described previously [31].

The TCGA database was analyzed using Kaplan-Meier plotter and GEPIA for NABP2 prognostic values in HCC, and the R "survival" package (version 3.5.2) was conducted to validate the prognostic value of NABP2 for HCC from the ICGC database. For the GEPIA and ICGC databases, the Kaplan-Meier curves were split at the median, and for the Kaplan-Meier plotter, the Kaplan-Meier curves were split at the best cutoff.

Clinical and pathological characteristics associated with NABP2 expression in HCC

Analyses of NABP2 expression and clinical characteristics in patients with HCC were conducted using the R "ggpubr" package, followed by validation via UALCAN (http://ualcan.path.uab.edu). Moreover, the major difference in NABP2 between the two experimental groups was analyzed with a *t*-test, and comparisons with more than two groups were conducted with an ANOVA.

Assays of cell viability, apoptosis, the cell cycle, and cell migration

Huh7 and LM3 cells in the exponential growth phase were grown for 6 h and transfected with NABP2 siRNA using Lipofectamine 2000 (Thermo Fisher Scientific, USA). The final concentration of siRNA was determined at 100 nM. The sequence of NABP2 siRNA1 was "GAACCAGAATGGGAATGGA", and siRNA2 was "CCCTGTTAGTAACGGCAAA". As previously described, Western blotting was used to detect NABP2 levels in HCC cells. This experiment was repeated three times. The primary antibodies were anti-NABP2 (Affinity Biosciences, DF13202, 1:1000) and anti-Tubbulin β (Bioworld, BS1482MH, 1:10000) [9].

Cell proliferation and viability were detected with a CCK8 assay (Vazyme; A31101AA), as previously described [32]. In conjunction with flow-cytometry analysis, cell cycle and cell apoptosis assays were performed as previously described [33]. Cell migration was detected using wound healing and a transwell migration assay, as described previously [34].

Consensus clustering analysis

Using the "ConsensusClusterPlus" package of R programming language, a hierarchical clustering algorithm was used to categorize HCC patients into the optimal number of clusters based on the expression of NABP2-related intersection genes [35].

LinkedOmics database analyses

NABP2 co-expression genes were statistically analyzed using Pearson's correlation coefficient. KEGG pathway, Gene Ontology biological process (GO_BP) analysis, transcription factors (TF), and kinase-target enrichment were obtained with gene set enrichment analysis (GSEA) in LinkedOmics functional modules.

GSEAs

Based on correlations between NABP2 expression and GSEAs, a comparison of different pathways between high- and low-NABP2 groups was performed. As the gene set database, we selected c2.cp.kegg.v7.0.symbols. Normalized enrichment scores (NESs) and *P* values were used to identify pathways enriched in each phenotype. The pathways with the following criteria were regarded as significantly enriched: q-value and normalized enrichment score (NES) > 1correspond to a false discovery rate of 0.05, respectively.

WGCNA analysis

Weightedgene co-expression Network Analysis (WGCNA) was used to construct weighted gene co-expression networks. The R package "WGCNA" was used to analyze 20% of the genes in TCGA, ICGA, and GEO datasets with the largest variance. The power (β) = 10 (R² = 0.98) in TCGA, power (β) = 14 (R² = 0.86) in ICGC and power (β) = 12 (R² = 0.91) in GEO were selected to build a co-expression network. We then selected module 0.2 to merge similar modules.

Model construction and validation for clinical prognoses

A potential independent risk factor for HCC survival was examined using univariate and multivariate analyses of data from TCGA, GSE14520, and ICGC databases. Moreover, the clinical application value of the risk score model was also determined by calculating the net benefits at each risk threshold by applying a decision curve analysis (DCA) [36]. With the help of the "regplot" package, a nomogram integrating risk signature, stage, age, sex, and grade for survival prediction was developed. Using the nomogram, a calibration curve was evaluated for predictive accuracy.

Nomogram construction and assessment

To measure risk and predict clinical outcomes in HCC patients, the nomogram was constructed with the "regplot" R package based on the independent clinical factors and risk scores. Calibration plots were used to evaluate the ability of the nomogram to predict HCC survival rates over the course of one, three, and five years. Nomogram clinical practicability was assessed using DCA.

Correlations of NABP2 expression and immune cells

For HCC, the CIBERSORT and Tumor Immune Estimation Resource (TIMER) databases were applied to study the association between NABP2 gene expression and immune cell infiltration. Then, a correlation analysis was performed using the Gene Expression Profiling Interactive Analysis (GEPIA) and TIMER databases to identify the correlation between NABP2 expression and immune marker genes.

Drug response analysis

The "pRRophetic" package was used to screen the drug responses of patients with high or low risk. CellMiner was used to provide the drug sensitivity for NABP2-related signature genes. We downloaded immune therapy data from the TCIA database (https://tcia.at/home) and analyzed the correlation between risk scores and immune therapy by applying R "ggpubr".

Statistical analysis

Statistical analysis was acquired with the TCGA and was determined with R-3.6.1. "Limma" and "vioplot" packages were applied to visualize the differences in NABP2 expression among the TCGA, ICGC, and CPTAC cohorts. ICGC's survival analysis was conducted using Kaplan-Meier Plotter. Spearman's correlation analysis between NABP2 gene expression in the GEPIA databases and TIMER was used to assess the association with NABP2 gene expression.

Results

The expression of NABP2 was upregulated in HCC

NABP2 expression was increased in HCC compared to normal tissues using data from the TCGA (Figure 1A), ICGC (Figure 1B), GCPIA (Figure 1C and 1D), and GEO (Figure 1E) datasets. Meanwhile, there was a significant increase in NABP2 protein expression in HCC using the HPA database (Figure 1F) and CPTAC database (Figure 1G). Moreover, HCC patient with the low protein expression level of NABP2 had a higher survival rate of (p=0.049) (Figure 1H). Furthermore, we detected NABP2 expression in 20 pairs of HCC tissue and compared normal tissue by using IHC. As shown in Figure 1I, NABP2 expression increased significantly in HCC compared to adjacent normal tissue. Then, according to Oncomine analyses, a higher level of NABP2 expression was observed in cancer tissues than in normal counterparts in the kidney, lung, and liver, while it was significantly lower in pancreatic cancers and brain cancers compared with the corresponding normal tissues (Figure S1). Moreover, several cancers were found to express NABP2 mRNA more strongly in the GEPIA database (<u>Figure S1</u>). The inconsistent data on NABP2 expression from Oncomine and GEPIA databases might be attributed to heterogeneity in the cancer populations of the two databases.

Prognostic role of NABP2 expression in HCC

GEPIA data revealed the prognostic role of NABP2 in HCC. We found that an increased expression of NABP2 in HCC was correlated with poor disease-free survival (DFS) and overall survival (OS) of HCC patients in the GEPIA database (Figure 2A). Subsequently, the ICGC database showed that the high expression of NABP2 was associated with poor OS among patients with HCC (Figure 2B). Furthermore, the Kaplan-Meier Plotter database analysis indicated high levels of NABP2 resulted in lower OS, Progression Free Survival (PFS), Relapse Free Survival (RFS), and Disease Specific Survival (DSS) rates (Figure 2C). To further reveal the prognostic value of NABP2, we performed univariate and multivariate Cox regression analyses using the HCC data from the TCGA and ICGC databases (Figure 2D and 2E). The results showed that NABP2 was an independent prognostic factor for OS in HCC. In conclusion, these results implied that NABP2 could serve as a potential prognostic indicator for HCC.

The correlation of NABP2 expression levels with the regulation of the cell cycle signal pathway and infiltrating immune cells

Next, GSEA analysis was performed to reveal the pathways regulated by NABP2. Figure S2A shows Alzheimer's disease and cell cycle abnormalities related to NABP2. On the contrary, changes in fatty acid metabolism and the peroxisome proliferator-activated receptor (PPAR) signaling pathway were observed to be associated with low expression of NABP2 (Figure S2B).

NABP2 expression in HCC was correlated with the number of infiltrating immune cells according to a TIMER database analysis (Figure S3). Figure S3A shows that NABP2 expression levels were associated with CD8⁺ T cell levels and the infiltration levels of B cells in HCC tissues. Next, we used CIBERSORT to predict immune cell infiltration in HCC using the TCGA dataset. Correlations between NABP2 expression and



Figure 1. Expression levels of NABP2 in hepatocellular carcinoma (HCC). A comparison of NABP2 expression between HCC tissues and normal tissues via the TCGA (A) and ICGC databases (B). A comparison of NABP2 expression in HCC tissues and normal tissues via the GCPIA database (C, D) and GTEx data (E). NABP2 protein expression in (F) the HPA database and (G) the CPTAC database. (H) The correlation between protein levels of NABP2 with the overall survival ratio of HCC patients. (I) Protein levels of NABP2 in HCC using IHC from xiangya hospital. The t test was to analyze the prominent of difference in NABP2 between 2 experimental groups.





Figure 2. NABP2 expression predicted the prognosis of hepatocellular carcinoma (HCC). (A) A correlation analysis of NABP2 and overall survival (OS), disease-free survival (DFS) in HCC using the Gene Expression Profiling Interactive Analysis database. (B) Correlation between NABP2 expression and OS from the International Cancer Genome Consortium (ICGC) database. (C) The correlation between NABP2 expression and OS, PFS, RFS, and DSS in HCC using the Kaplan-Meier Plotter database. Cox regression and ROC curve analyses showed associations between NABP2 expression and clinical factors in HCC survival in (D) the Cancer Genome Atlas database and (E) the ICGC database.

the infiltration levels of neutrophils and eosinophils were observed in HCC (Figure S3B). Furthermore, we found that CD4 memory resting T cells and naive B cells were significantly reduced in the group with high expression of NABP2 (Figure S3C). The results indicated that NABP2 regulated immune cell infiltration in HCC.

NABP2 is an oncogene in HCC

To assess the role of NABP2 in HCC cells, a siRNA targeting NABP2 was transfected into LM3 and Huh7 cells. The results showed that silencing of NABP2 inhibited cell proliferation (Figure 3A and 3B), induced apoptosis (Figure 3C), and blocked the cell cycle in S phase in LM3 cells, and blocked the cell cycle in GO/G1 phase in Huh7 cells (Figure 3D). Moreover, transwell and wound-healing assays showed that silencing NABP2 inhibited HCC cell migration (Figure 3E and 3F). These results indicated that NABP2 promoted HCC progression by increasing cell viability and repressing cell apoptosis in HCC.

To further reveal the role of NABP2 in immunoregulation in HCC, we explored the expression of chemokines and cytokines in NABP2silenced HCC cells. As shown in **Figure 3G**, IL4, IL9, and IL23 were significantly decreased in LM3 and Huh7 cells by silencing NABP2. IL17B was down-regulated by silencing NABP2 in Huh7 cells (IL17B expression was not detected in LM3 cells). Together, according to these results, NABP2 could play a critical role in immunoregulation by regulating IL4, IL9, IL23, and IL17B expression in HCC cells.

Identifying hub modules and enrichment analysis

To further reveal the mechanisms of NABP2associated regulations, we analyzed the NABP2-associated core module in HCC using WGCNA datasets. As shown in **Figure 4A**, seven, six modules, and eight modules were obtained for HCC from the GEO, ICGC, and TCGA databases, respectively. The turquoise module was statistically associated with NABP2 in the TCGA cohort dataset, the grey module was statistically correlated with NABP2 in the ICGC cohort dataset, and the yellow module was statistically associated with NABP2 in the GEO cohort dataset, as shown in **Figure 4A**. Considering that the focus of our study was NABP2, 668 hub genes were obtained in turquoise modules from the TCGA-LIHC dataset, 2,995 hub genes were obtained in grey modules from the ICGC dataset, and 169 hub genes were obtained in yellow modules from the GEO dataset. Subsequently, a Venn diagram illustrated 45 overlapped NABP2-related genes from the TCGA, ICGC, and GEO cohort datasets (**Figure 4B**). Moreover, GO and KEGG pathway analyses showed that 45 NABP2-related common genes were related to the p53 signaling pathway, cell cycle, and DNA replication (**Figure 4C** and **4D**).

NABP2-related gene cluster biological characteristics

Subsequently, an unsupervised consensus analysis was conducted based on the 45 NABP2-related genes. Two subtypes of HCC were identified, and the patients in cluster A showed poorer prognoses than those in cluster B (**Figure 5A-C**).

To further explore the heterogeneity of two clusters, differential expression and immune cellinfiltration analyses were performed using the TCGA, ICGC, and GEO databases (**Figure 6**). The 45 NABP2-related genes were more prevalent in cluster A than in cluster B (**Figure 6**). Moreover, the activated CD4 T cells and T helper cells were increased in cluster A, while eosinophils were decreased in cluster A in the TCGA, ICGC, and GEO databases (**Figure 6**).

Construction and verification of a NABP2-related gene risk signature

We analyzed the 440 differentially expressed genes between two clusters that overlapped among the three databases (**Figure 7A**). The KEGG and GO enrichment analyses showed that the 440 genes were associated with DNA replication and the cell cycle (**Figure 7B** and **7C**). We then used the 440 cluster-related differential genes to establish a prognostic signature by employing the LASSO Cox regression analysis based on the TCGA datasets (**Figure 7D**). Considering the median of the risk scores, patients with HCC were then classified into lowrisk and high-risk groups.

Risk score = PSRC1 × 0.0789 + KIF20A × 0.0665 + GINS1 × 0.0057 + HDAC2 × 0.0052.





Figure 3. NABP2 siRNA influenced NABP2 expression, LM3 and Huh7 cell proliferation, apoptosis, cell cycle, and cell migration (n = 3). (A) Protein expression of NABP2 in Huh7 and LM3 cells. (B) Silencing NABP2 inhibited cell proliferation, (C) induced hepatocellular carcinoma cell apoptosis, (D) regulated the cell cycle, and (E and F) cell migration. (G) The expression of cytokines in NABP2 silenced Huh7 and LM3 cells. The mean \pm SD were used to summarize the results. Student's *t*-tests were used to analyze comparisons between groups. Statistical significance: *P < 0.05, **P < 0.001, ***P < 0.001, ****P < 0.0001.



Figure 4. Enrichment analysis in hepatocellular carcinoma patients. (A) An analysis of weighted gene co-expression networks was conducted to generate networks of co-expressed genes according to the Cancer Genome Atlas (TCGA) database, International Cancer Genome Consortium (ICGC) dataset, and the GEO database. (B) Venn diagram of the intersection of NABP2-related genes from the TCGA database, ICGC dataset, and GEO database. GO (C) and KEGG (D) analyses of NABP2-related intersecting genes.



Figure 5. NABP2 intersection gene cluster analysis. Two distinct clusters, consensus clustering cumulative distribution function (CDF), relative change in the area under the CDF curve, and Kaplan-Meier curve survival analysis in the Cancer Genome Atlas (A), International Cancer Genome Consortium (B) and GEO (C) cohort datasets.





Figure 6. Correlation analysis of gene clusters with prognosis-related genes and immune cell-infiltration. Based on the Cancer Genome Atlas (A), International Cancer Genome Consortium (B), and GEO database (C), two clusters and genes related to immune cell infiltration were compared. A comparison of two clusters and genes associated with prognosis and immune cell infiltration was conducted.



Figure 7. Construction of a NABP2-related risk signature for hepatocellular carcinoma. (A) An intersection gene cluster diagram displaying NABP2-related genes from the Cancer Genome Atlas (TCGA), International Cancer Genome Consortium (ICGC), and GSE14520 datasets. The cluster plot of the GO (B) and KEGG (C) intersection genes from TCGA, ICGC, and GSE14520 datasets. (D) LASSO Cox regression analysis. (E) Survival analysis. (F) t-SNE analysis. (G) ROC curve. (H) Distribution of the risk scores and heat map of NABP2-related signature genes.

Based on these results, we found that the prognoses were worse in the high-risk group than in the low-risk group (**Figure 7E**). A t-SNE analysis showed a distinct distribution of genes in the high- and low-risk groups (**Figure 7F**), and the AUC was 0.766 for 1-year, 0.696 for 2-year, and 0.671 for 3-year survival (**Figure 7G**). As shown in **Figure 7H**, the survival rate was decreased and the expression levels of risk genes (i.e., PCRC1, HDAC2, KIF20A, and GINS1) were increased with increases in the risk score.

Next, the ICGC and GEO datasets were used to verify the accuracy of the NABP2-related risk signature in HCC. We found that the high-risk group had a significantly higher death rate than the low-risk group in both the ICGC and GEO datasets (**Figure 8A** and **8E**). A t-SNE analysis showed a distinct distribution between the high- and low-risk groups (**Figure 8B** and **8F**). The AUC was 0.724 and 0.622 for 1-year, 0.700 and 0.638 for 2-year, and 0.735 and 0.623 for 3-year survival, respectively (**Figure 8C** and **8G**). As shown in **Figure 8D** and **8H**, the HCC patients with high risk had a poor survival rate and increased expression of PCRC1, HDAC2, KIF2OA, and GINS1.

Prognostic value of the NABP2-related risk signature

Next, we analyzed the prognostic role of the NABP2-related risk signature. A Cox regression analysis showed that risk score and stage were independently associated with better outcomes in HCC in both the TCGA (**Figure 9A**) and ICGC databases (**Figure 9C**) and GEO database (**Figure 9E**), respectively. To estimate the clinical utility of the risk signature, ROC curves indicated that the AUCs of stages were 0.638, 0.814, and 0.691 using the TCGA database (**Figure 9B**), ICGC database (**Figure 9D**), and GEO database (**Figure 9B**), ICGC database (**Figure 9D**), and GEO database (**Figure 9F**), respectively. The results indicated that NABP2-related risk was a highly reliable signature.

According to a gene set variation analysis (GSVA), NABP2-related genes were primarily enriched in the cell cycle, p53 signaling pathway, and pyrimidine metabolism (Figure S4A). A

GO analysis showed that NABP2-related genes had a strong association with tubulin binding, nuclear division, spindle, chromosomal region, organelle fission, and microtubule binding vital biological processes (Figure S4B). Results from a KEGG analysis showed that NABP2-related genes were primarily associated with DNA replication, regulation of the mitotic cell cycle phase transition, and humoral immune responses (Figure S4C).

Nomogram construction and prediction evaluation

Based on independent prognostic factors, we developed a nomogram to predict outcomes in HCC over one, three, and five years (**Figure 10**). Nomogram scores were calculated by adding the points from all variables. Then, HCC patient survival rates were predicted by drawing a vertical line from the total points to the survival prediction axis. In the calibration plot, nomograms with 1-, 3-, and 5-year predictions indicated that the nomograms performed well. To further assess the net clinical benefits of multiple prognostic factors, we used a DCA. According to the concordance index curves, risk score was a more accurate predictor of HCC survival when compared to other clinical factors.

Estimation of the immune microenvironment and tumor mutational burden according to the risk signature

To confirm the accuracy of the cluster scheme described above, we compared the immune infiltration and immune-related pathways between the high and low groups. Based on the gene expression profiles, we determined the stromal score, estimate score, tumor purity, and immune score for each HCC sample using the ESTIMATE algorithm. Compared to the lowrisk group, the high-risk group had significantly lower stromal scores and significantly higher estimate scores (**Figure 11A**). Based on these results, it was clear that these two groups had completely different characteristics in terms of immune cell infiltration and immune-related functions. Th2_cells, aDCs, macrophages, and



Figure 8. NABP2-related gene prognostic signature of hepatocellular carcinoma (HCC) in the International Cancer Genome Consortium (ICGC) dataset and verified in the GEO dataset. (A) Survival analysis, (B) t-SNE analysis, and (C) ROC curve of HCC patients with low- or high-risk scores using the ICGC dataset. (D) Survival analysis, (E) t-SNE analysis, and (F) ROC curve of HCC patients with high- or low-risk scores using the GEO dataset. Risk score distribution and signature gene heatmap for CRGs using the Cancer Genome Atlas dataset (G) and GEO dataset (H).

Treg cells were enriched in the high-risk group, but cytolytic_activity, mast_cells, neutrophils, NK_cells and type_II_IFN_responses were enriched in the low-risk group (**Figure 11B**). As shown in **Figure 11C**, aDCs and MHC_class_I had a positive and significant correlation with



Figure 9. Prognostic role of the NABP2-related risk signature and clinicopathological characteristics in hepatocellular carcinoma. The forest plots for univariate and multivariate Cox regression analyses using the Cancer Genome Atlas (TCGA) databases (A), International Cancer Genome Consortium (ICGC) databases (C), and GEO databases (E), and the areas under the ROC curve regarding gender, age, risk score, and stage using TCGA databases (B), ICGC databases (D), and GEO databases (F).



Figure 10. Construction of the prediction model and evaluation of the predicted outcomes using the Cancer Genome Atlas database (A), International Cancer Genome Consortium database (B) and GEO database (C).



Figure 11. Analysis of the relationship between the immune microenvironment and NABP2-related genes in hepatocellular carcinoma. Comparison of low-risk and high-risk groups regarding stromal scores, immune scores, and estimation scores (A), as well as immune-related functions (B). Correlation between low-risk and high-risk groups in terms of immune-related functions and NABP2-related genes (C). Boxplot illustrating the differences between chemokines, cytokines, and immune checkpoints (D).

the NABP2-related risk signature, while type_ II_IFN_responses, mast_cells, and NK_cells had a negative correlation with the NABP2related risk signature. Finally, we compared chemokines, cytokines, and immune checkpoint expressions between the groups (**Figure 11D**). Chemokines CCL3, CXCL6, CCL20, CXCL8, CCL4L2, CCL28, CCL13, CXCL1, CX3CL1, CXCL10, CXCL3, CXCL13, CXCL1, CX3CL1, CXCL10, CXCL3, CXCL13, CXCL5, CCL4, CCL3L1, CCL26, CXCL11, and CXCL9 exhibited high expression levels and chemokines CCL14, CCL16, CXCL2, and CCL21 exhibited low expression levels in the high-risk group. A high level of expression for some cytokine genes (IL23A, IL7, IL18, IL12A, IL10,



IL17D, IL15, and IL1B) was detected in highrisk individuals, while high expression levels of IL27, IL32, and IL32 were observed in lowrisk individuals. Furthermore, there was high expression of immune checkpoints (CD200R1, CD160, TNFRSF14, CTLA4, PDCD1, and others) in high-risk individuals.

We next analyzed the mutation rate and tumor mutation burden (TMB) between the high- and low-risk groups. Waterfall plots showed the mutation distribution of the top 20 most frequently mutated genes in HCC patients with low or high risk (Figure S5A). Patients in the high-risk group had a higher mutation frequency. The patients with high risk had a significantly higher mutation frequency for TP53 than patients with low risk, while the mutation frequency of AXIN1 in the low-risk group was comparable to that in the high-risk group. A TMB analysis indicated that no significant TMB difference was observed between the high and low groups, but the H-TMB with high-risk group had poor survival probability (<u>Figure S5B</u>).

Predictions for HCC drug treatments

To test if NABP2 could be used to treat HCC patients individually, we compared the IC_{50} values of different chemotherapy agents between NABP2-related risk signature groups. Drug susceptibility results showed that patients with high risk had lower IC_{50} values for A-443654, BI-2536, gemcitabine, GW843682X, JW-7-52-1, paclitaxel, rapamycin, and vinorelbine than patients with low risk, implying that the above therapy may be more beneficial to these patients (**Figure 12**).

Discussion

With high morbidity and mortality, HCC is a growing burden on global health [11]. Despite the approval of new drugs for HCC treatment, the effectiveness in advanced HCC patients remains unsatisfactory and the 5-year survival rate remains low [12]. Hence, in order to diagnose and treat HCC as early as possible, new targets are urgently needed. To our knowledge, this is the first study in which NABP2 has been identified as an oncogene in HCC patients and a NABP2-related prognosis signature was described.

Being a regulator of DNA replication, recombination, and repair, NABP2 has recently been identified as related to cancer susceptibility in mice [13]. Through bioinformatics and IHC analysis, we found that higher levels of NABP2 mRNA and protein were notably associated with poorer DSS, OS, PFS, and RFS of patients with HCC. Therefore, NABP2 could function as a prognostic biomarker of HCC. During stalled DNA replication forks, NABP2 functions by binding to ssDNA to facilitate the repair of doublestrand DNA breaks and oxidized DNA adducts [8, 14-18]. A recent study showed that NABP2 might be related to the integrator complex, RNA polymerase II, and transcriptional termination factors to promote mRNA transcriptional termination [19]. Also, NABP2 could protect p53 against ubiquitin-mediated degradation in DNA damage events and activate p53-mediated expression of p21 [10], which was consistent with our findings. In our study, pathways such as the cell cycle, p53 signaling pathway, bladder cancer, pyrimidine metabolism, and glutathione metabolism were enriched in cells with NABP2 high expression in HCC. As a further evidence for NABP2's tumor-promoting role in HCC, we found that silencing NABP2 suppressed cell apoptosis, cell proliferation, and cell migration in HCC cell lines. Recently, evidence has been growing that suggests tumor suppressor genes or oncogenes can modulate the immune system to inhibit or enhance tumor growth [20, 21]. In our research, there was a correlation between NABP2 expression and the infiltration of neutrophils, B cells, CD4⁺ cells, CD8⁺ TDCs, and macrophages, which implied that NABP2 participated in regulating cancer immunity. NABP2 was also related to Th2 marker genes (STAT6 and STAT5A), which contribute markedly to the infiltration of T helper cells. Our study further showed a strong relationship between NABP2 and T cell exhaustion. CTLA4 is an inhibitory T cell receptor and negatively regulated the survival, proliferation, and cytokine production of T cells [22, 23]. Higher NABP2 expression levels were associated with more exhausted T cell markers (CTLA4 and LAG3) and inhibitory immune checkpoint proteins. The progression of various types of cancer is influenced by chemokines and cytokines by recruiting immune cells to the tumor [24]. Our study identified a positive correlation between NABP2 expressions and IL-4, IL-9, IL-17B, and IL23 levels, while silencing NABP2 drastically reduced expression of these cytokines in HCC cells. Thus, we hypothesized that over-expression of NABP2 was involved in HCC progression by regulating the p53 signaling pathway, cell cycle process, and recruitment of immune cells.

Accumulating evidence has shown that gene expression signals are precise predictors of the prognosis in HCC patients. Huang et al. developed and validated a prognostic signature based on eight genes related to NOL12 [25]. This study provided additional evidence that prognostic biomarkers and therapeutic targets can be identified in HCC based on a prognostic signature [25]. Given our results, which showed



Figure 12. Drug sensitivity analysis of NABP2-related genes and risk scores showed an association with chemotherapeutic sensitivity in hepatocellular carcinoma. A. Differential chemotherapeutic responses in high- or low-risk patients according to the gene signature model. B. Relationship between drug sensitivity and risk scores according to the gene signature model.

that NABP2 had the potential to be a valuable diagnostic and prognostic biomarker, the NABP2-related gene signature was capable of predicting patient outcomes with accuracy. We obtained 440 cluster-related differential genes that were both related to the cell cycle and DNA replication. Then, based on screening four NABP2-related genes, we constructed a NABP2-related gene risk signature. The risk signature showed excellent performance with a high level of accuracy in predicting 1-, 2-, and 3-year survival rates. In addition, the risk signature was independently determined through independent prognostic analysis. As confirmed by the DCA, ROC curve, and calibration curve, our nomogram could provide clinicians with an effective, convenient, and practical tool for predicting prognoses in HCC patients. Immune cell infiltration in the tumor microenvironment (TME) critically influenced tumor metastasis, cell survival, and resistance to therapy [26-29]. Our study indicated that the TME infiltration characteristics, immune cell functions, and expression of chemokines and cytokines were different between high-risk and low-risk patients. Moreover, high-risk patients expressed higher levels of immune checkpoint molecules, indicating that they may benefit more from immune checkpoint inhibitor treatment.

To some extent, drug repurposing can identify "old drugs for new use", which drives drug development [30]. The correlation of a NABP2related signature with the efficacy of drugs was analyzed using the TCGA-LIHC dataset. A total of eight potential chemicals with a lower IC_{50} had a significant association with high-risk NABP2. These results indicated that the eight drugs might be candidates for treating HCC patients with high-risk NABP2.

Conclusions

In conclusion, our study provided comprehensive and strong evidence suggesting that NABP2 can serve as a potential prognostic and therapeutic biomarker for hepatocarcinogenesis. Additionally, the NABP2-related risk signature can function as a prognostic factor in HCC treatment.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (8170-3149, 32170529), the Innovation-Driven Project of Central South University (2020CX044), the Scientific Research Fund of Hunan Provincial Education Department (21A0614), the Technology Innovation Project of Huaihua (2021R3102), the Scientific Research Cultivating Fund of Hunan University of Medicine (2020-5), the Innovation and Entrepreneurship Training Program for Hunan College Students (S2022122140027), the National Natural Sciences Foundation of Hunan province (2023JJ50438 and 2022JJ40805), and the National Natural Science Foundation of China (82103638).

Disclosure of conflict of interest

None.

Address correspondence to: Xiaozhen Peng and Yiya Zhang, Department of Dermatology, Xiangya Hospital, Central South University, Changsha, Hunan, China. Tel: +86-18374589060; E-mail: peng112112@163.com (XZP); Tel: +86-156751-74917; E-mail: yiya0108@csu.edu.cn (YYZ)

References

- [1] Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin DM, Forman D and Bray F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015; 136: E359-386.
- [2] Heimbach JK, Kulik LM, Finn RS, Sirlin CB, Abecassis MM, Roberts LR, Zhu AX, Murad MH and Marrero JA. AASLD guidelines for the treatment of hepatocellular carcinoma. Hepatology 2018; 67: 358-380.
- [3] Giannini EG, Farinati F, Ciccarese F, Pecorelli A, Rapaccini GL, Di Marco M, Benvegnu L, Caturelli E, Zoli M, Borzio F, Chiaramonte M and Trevisani F; Italian Liver Cancer (ITA.LI.CA) group. Prognosis of untreated hepatocellular carcinoma. Hepatology 2015; 61: 184-190.
- [4] Ge PL, Li SF, Wang WW, Li CB, Fu YB, Feng ZK, Li L, Zhang G, Gao ZQ, Dang XW and Wu Y. Prognostic values of immune scores and im-

mune microenvironment-related genes for hepatocellular carcinoma. Aging (Albany NY) 2020; 12: 5479-5499.

- [5] Wold MS. Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. Annu Rev Biochem 1997; 66: 61-92.
- [6] Bhatti S, Kozlov S, Farooqi AA, Naqi A, Lavin M and Khanna KK. ATM protein kinase: the linchpin of cellular defenses to stress. Cell Mol Life Sci 2011; 68: 2977-3006.
- [7] Helleday T, Lo J, van Gent DC and Engelward BP. DNA double-strand break repair: from mechanistic understanding to cancer treatment. DNA Repair (Amst) 2007; 6: 923-935.
- [8] Richard DJ, Bolderson E, Cubeddu L, Wadsworth RI, Savage K, Sharma GG, Nicolette ML, Tsvetanov S, McIlwraith MJ, Pandita RK, Takeda S, Hay RT, Gautier J, West SC, Paull TT, Pandita TK, White MF and Khanna KK. Singlestranded DNA-binding protein hSSB1 is critical for genomic stability. Nature 2008; 453: 677-681.
- [9] Xu S, Feng Z, Zhang M, Wu Y, Sang Y, Xu H, Lv X, Hu K, Cao J, Zhang R, Chen L, Liu M, Yun JP, Zeng YX and Kang T. hSSB1 binds and protects p21 from ubiquitin-mediated degradation and positively correlates with p21 in human hepatocellular carcinomas. Oncogene 2011; 30: 2219-2229.
- [10] Xu S, Wu Y, Chen Q, Cao J, Hu K, Tang J, Sang Y, Lai F, Wang L, Zhang R, Li SP, Zeng YX, Yin Y and Kang T. hSSB1 regulates both the stability and the transcriptional activity of p53. Cell Res 2013; 23: 423-435.
- [11] Zhou M, Zhang Q, Zhao J, Liao M, Wen S and Yang M. Phosphorylation of Bcl-2 plays an important role in glycochenodeoxycholate-induced survival and chemoresistance in HCC. Oncol Rep 2017; 38: 1742-1750.
- [12] Bangaru S, Marrero JA and Singal AG. Review article: new therapeutic interventions for advanced hepatocellular carcinoma. Aliment Pharmacol Ther 2020; 51: 78-89.
- [13] Shi W, Bain AL, Schwer B, Al-Ejeh F, Smith C, Wong L, Chai H, Miranda MS, Ho U, Kawaguchi M, Miura Y, Finnie JW, Wall M, Heierhorst J, Wicking C, Spring KJ, Alt FW and Khanna KK. Essential developmental, genomic stability, and tumour suppressor functions of the mouse orthologue of hSSB1/NABP2. PLoS Genet 2013; 9: e1003298.
- [14] Ashton NW, Paquet N, Shirran SL, Bolderson E, Kariawasam R, Touma C, Fallahbaghery A, Gamsjaeger R, Cubeddu L, Botting C, Pollock PM, O'Byrne KJ and Richard DJ. hSSB1 phosphorylation is dynamically regulated by DNA-PK and PPP-family protein phosphatases. DNA Repair (Amst) 2017; 54: 30-39.

- [15] Bolderson E, Petermann E, Croft L, Suraweera A, Pandita RK, Pandita TK, Helleday T, Khanna KK and Richard DJ. Human single-stranded DNA binding protein 1 (hSSB1/NABP2) is required for the stability and repair of stalled replication forks. Nucleic Acids Res 2014; 42: 6326-6336.
- [16] Paquet N, Adams MN, Leong V, Ashton NW, Touma C, Gamsjaeger R, Cubeddu L, Beard S, Burgess JT, Bolderson E, O'Byrne KJ and Richard DJ. hSSB1 (NABP2/OBFC2B) is required for the repair of 8-oxo-guanine by the hOGG1mediated base excision repair pathway. Nucleic Acids Res 2015; 43: 8817-8829.
- [17] Richard DJ, Cubeddu L, Urquhart AJ, Bain A, Bolderson E, Menon D, White MF and Khanna KK. hSSB1 interacts directly with the MRN complex stimulating its recruitment to DNA double-strand breaks and its endo-nuclease activity. Nucleic Acids Res 2011; 39: 3643-3651.
- [18] Richard DJ, Savage K, Bolderson E, Cubeddu L, So S, Ghita M, Chen DJ, White MF, Richard K, Prise KM, Schettino G and Khanna KK. hSSB1 rapidly binds at the sites of DNA double-strand breaks and is required for the efficient recruitment of the MRN complex. Nucleic Acids Res 2011; 39: 1692-1702.
- [19] Skaar JR, Ferris AL, Wu X, Saraf A, Khanna KK, Florens L, Washburn MP, Hughes SH and Pagano M. The integrator complex controls the termination of transcription at diverse classes of gene targets. Cell Res 2015; 25: 288-305.
- [20] Bi KW, Wei XG, Qin XX and Li B. BTK has potential to be a prognostic factor for lung adenocarcinoma and an indicator for tumor microenvironment remodeling: a study based on TCGA data mining. Front Oncol 2020; 10: 424.
- [21] Chen J, Wang Z, Wang W, Ren S, Xue J, Zhong L, Jiang T, Wei H and Zhang C. SYT16 is a prognostic biomarker and correlated with immune infiltrates in glioma: a study based on TCGA data. Int Immunopharmacol 2020; 84: 106490.
- [22] Yang F, Li Y, Zhang Q, Tan L, Peng L and Zhao Y. The effect of immunosuppressive drugs on MDSCs in transplantation. J Immunol Res 2018; 2018: 5414808.
- [23] Ouyang Z, Wu H, Li L, Luo Y, Li X and Huang G. Regulatory T cells in the immunotherapy of melanoma. Tumour Biol 2016; 37: 77-85.
- [24] Malik A and Kanneganti TD. Function and regulation of IL-1alpha in inflammatory diseases and cancer. Immunol Rev 2018; 281: 124-137.
- [25] Huang J, Kang W, Pan S, Yu C, Jie Z and Chen C. NOL12 as an oncogenic biomarker promotes hepatocellular carcinoma growth and metastasis. Oxid Med Cell Longev 2022; 2022: 6891155.

- [26] Joyce JA and Fearon DT. T cell exclusion, immune privilege, and the tumor microenvironment. Science 2015; 348: 74-80.
- [27] Waniczek D, Lorenc Z, Snietura M, Wesecki M, Kopec A and Muc-Wierzgon M. Tumor-associated macrophages and regulatory T cells infiltration and the clinical outcome in colorectal cancer. Arch Immunol Ther Exp (Warsz) 2017; 65: 445-454.
- [28] Noy R and Pollard JW. Tumor-associated macrophages: from mechanisms to therapy. Immunity 2014; 41: 49-61.
- [29] Sato E, Olson SH, Ahn J, Bundy B, Nishikawa H, Qian F, Jungbluth AA, Frosina D, Gnjatic S, Ambrosone C, Kepner J, Odunsi T, Ritter G, Lele S, Chen YT, Ohtani H, Old LJ and Odunsi K. Intraepithelial CD8+ tumor-infiltrating lymphocytes and a high CD8+/regulatory T cell ratio are associated with favorable prognosis in ovarian cancer. Proc Natl Acad Sci U S A 2005; 102: 18538-18543.
- [30] Shah RR and Stonier PD. Repurposing old drugs in oncology: opportunities with clinical and regulatory challenges ahead. J Clin Pharm Ther 2019; 44: 6-22.
- [31] Lu S, Yu Z, Xiao Z and Zhang Y. Gene signatures and prognostic values of m(6)A genes in nasopharyngeal carcinoma. Front Oncol 2020; 10: 875.
- [32] Tiemin P, Fanzheng M, Peng X, Jihua H, Ruipeng S, Yaliang L, Yan W, Junlin X, Qingfu L, Zhefeng H, Jian L, Zihao G, Guoxing L, Boshi S, Ming Z, Qinghui M, Desen L and Lianxin L. MUC13 promotes intrahepatic cholangiocarcinoma progression via EGFR/PI3K/AKT pathways. J Hepatol 2020; 72: 761-773.

- [33] Wang XK, Liao XW, Yang CK, Liu ZQ, Han QF, Zhou X, Zhang LB, Deng T, Gong YZ, Huang JL, Huang R, Han CY, Yu TD, Su H, Ye XP, Peng T and Zhu GZ. Oncogene PLCE1 may be a diagnostic biomarker and prognostic biomarker by influencing cell cycle, proliferation, migration, and invasion ability in hepatocellular carcinoma cell lines. J Cell Physiol 2020; 235: 7003-7017.
- [34] Liu Z, Yu Y, Huang Z, Kong Y, Hu X, Xiao W, Quan J and Fan X. CircRNA-5692 inhibits the progression of hepatocellular carcinoma by sponging miR-328-5p to enhance DAB2IP expression. Cell Death Dis 2019; 10: 900.
- [35] Wilkerson MD and Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence assessments and item tracking. Bioinformatics 2010; 26: 1572-1573.
- [36] Van Calster B, Wynants L, Verbeek JFM, Verbakel JY, Christodoulou E, Vickers AJ, Roobol MJ and Steyerberg EW. Reporting and interpreting decision curve analysis: a guide for investigators. Eur Urol 2018; 74: 796-804.



Figure S1. The differential NABP2 expression between multi-tumor tissues and normal tissues using the Oncomine database and GEPIA database.





Figure S2. Analyses of GSEA indicated pathways enriched in high expression phenotype (A) and low expression phenotype of NABP2 (B).



Figure S3. Analysis of tumor immune cell infiltration relative to NABP2 expression in HCC. A. Correlation of immune cells with NABP2 expression using the TIMER database. B. The relationship between immune cell infiltration from TCGA using CIBERSORT and NABP2 expression identified by Pearson test. C. The infiltration levels of immune cells with high/low expressed NABP2 in HCC.



Figure S4. Enrichment analysis GSVA analysis (A), GO pathways of NABP2-related genes on biological progress, the cellular component and molecular functions (B), and KEGG analysis (C) of NABP2-related genes on biological process.



Figure S5. Different NABP2-related risk signature groups for tumor mutation analysis. A. The waterfall plot of mutation analysis. B. A comparison of the tumor mutation burden of NABP2-related genes in HCC patients with low-/ high-risk.