Identification of biomarkers of hepatocellular carcinoma gene prognosis based on the immune-related IncRNA signature of transcriptome data

Andi Ma1 · Yukai Sun2 · Racheal O. Ogbodu3 · Ling Xiao4 · Haibin Deng5 · Hui Zhou6

Received: 7 February 2023 / Revised: 6 March 2023 / Accepted: 8 March 2023 / Published online: 28 March 2023
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2023

Abstract

Background Long non-coding RNAs (lncRNAs) are well established to have an important role in cancer. The goal of this research was to investigate the prognostic usefulness of putative immune-related lncRNAs in hepatocellular carcinoma (HCC).

Methods The developed lncRNA signature was validated using 343 HCC patients from The Cancer Genome Atlas (TCGA) and 81 samples from Gene Expression Omnibus (GEO). Cox regression and Least Absolute Shrinkage and Selection Operator (LASSO) analysis were used to analyze immune-related lncRNAs for HCC prognosis. Patients in the low-risk group survived substantially longer than those in the high-risk group ($P < 0.05$). The discovered signal might be a useful prognostic factor for predicting patient survival. Overall survival predicted some clinical net improvements, according to the nomogram. Numerous enrichment approaches (including gene set enrichment analysis) were utilized to investigate the underlying mechanisms.

Results Drug metabolism, mTOR, and p53 signaling pathways were associated with high-risk groups. When the expression of lncRNA PRRT3-AS1 was silenced in HepG2 cells, the proliferation, migration, and invasion abilities of HepG2 cells were decreased, and apoptosis was enhanced. In the supernatant from HepG2 cells with PRRT3-AS1 knockdown, the anti-inflammatory factors IL-10 and TGF-1 were induced, whereas the pro-inflammatory factors IL-1β, TNF-α, and IL-6 were reduced ($P < 0.05$). After PRRT3-AS1 knockdown, the protein expression of CD24, THY1, LYN, CD47, and TRAF2 in HepG2 cells was attenuated ($P < 0.05$).

Conclusion The discovery of five immune-related lncRNA signatures has significant therapeutic significance for predicting patient prognosis and directing personalized treatment for patients with HCC, which requires additional prospective confirmation.

Keywords Long non-coding RNAs · Hepatocellular carcinoma · Signature · Prognosis · Immune-related signature

Introduction

Hepatocellular carcinoma (HCC) is the sixth most common cancer in the world and the fourth leading cause of cancer-related deaths, and its annual incidence is increasing rapidly (Bray et al. 2018). Treatments for HCC include surgical resection, transplantation, ablation, transarterial chemoembolization, and tyrosine kinase inhibitors such as sorafenib, lenvatinib, and regorafenib, which have proven to be effective treatments for improving patient survival (Forner et al. 2018). The latest research unveiled that through immune checkpoint treatment, cytotoxic T lymphocyte antigen 4

---

1 Hunan University of Medicine School of Public Health and Laboratory Medicine, Huaibin 418000, China
2 Max Delbruck Centrum fur Molekulare Medizin Experimental and Clinical Research Center (MDC), AG Translational Oncology of Solid Tumors, Robert-Rössle-Straße 10, 13125 Berlin, Germany
3 Institute of Physiology, Charité-Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany
4 Department of Histology and Embryology, School of Basic Medical Sciences, Central South University Xiangya School of Medicine, Changsha 410013, China
5 Department of General Thoracic Surgery, Inselspital Universitätsklinik Bern, Murtenstrasse 50, 3008 Bern, Switzerland
6 Department of Lymphoma and Hematology, the Affiliated Cancer Hospital of Xiangya School of Medicine, Central South University, Hunan Cancer Hospital, Changsha 410013, Hunan, China
The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases. This expression profile aided the identification of potential immune-related lncRNA signals to predict the survival of patients with HCC.

We established a risk model based on immune-related lncRNAs to predict the survival of HCC patients. In addition, we predicted the regulatory pathways of lncRNAs using GSEA. Five lncRNAs involved in the immune system were identified in this work with the potential to be used in the diagnosis, prognosis, and therapy of HCC (Fig. S1). Thus, PRRT3-AS1 may be a potential target for HCC.

Material and methods

Patient datasets

The study employed the gene expression profiles and clinical data from TCGA (https://cancergenome.nih.gov/, January 06, 2020), enrolled 50 normal and 343 patients with HCC with survival data, and the GSE27150 dataset from the GEO database (https://www.ncbi.nlm.nih.gov/), including 81 patients with HCC.

Identification of DEGs

The immune-related gene set in the GSEA database (gene set: IMMUNE_RESPONSE, M19817, and IMMUNE_SYSTEM_PROCESS, M13664) was identified, and IRGs were obtained using genetic data from TCGA through the co-expression method (co-data moderated random forest (CoRF) > 0.4, P < 0.001). Then, using the intersection, 24 IRGs co-expressed in the TCGA and GEO databases were obtained. Subsequently, DEGs were identified using the “limma” package (http://www.bioconductor.org/) in R software version 3.6.2. The threshold value of DEGs was set as a log-fold change (FC) > 1.0 and P value < 0.05. Furthermore, the “pheatmap” package in R software was used for clustering analysis.

PPI network analysis

The PPI network was constructed based on online search databases (STRING; http://string-db.org) (version 11.0) (Franceschini et al. 2013) to identify interactive genes for making predictions. Analyzing functional interactions between proteins may provide insights into disease development or development mechanisms, including direct (physical) and indirect (functional) associations. In this study, the STRING database was used to evaluate their functional associations and construct a PPI network, and interactions with a combined score > 0.4 were considered statistically significant. Subsequently, the R software was employed to obtain a histogram of the node number.
Construction of the prognostic immune gene signature

Univariate and least absolute shrinkage and selection operator (lasso) Cox regression analyses (O’Connor 2000) were used to identify relevant immune genes to predict and construct prognostic gene signatures. \( P < 0.001 \) was considered statistically significant in univariate Cox regression analysis. The prognostic gene signature was calculated as follows: \( \text{risk score} = (\text{coefficient of gene}_1 \times \text{gene}_1 \text{ expression}) + (\text{coefficient of gene}_2 \times \text{gene}_2 \text{ expression}) + (\text{coefficient of gene}_N \times \text{Gene}_N \text{ expression}). \) The R package “survival” was used to explore the best-dividing line of patient risk score, and draw the Kaplan–Meier survival curve (\( P < 0.001 \)). Next, the “CorrPlot” R package was used to construct a correlation matrix between different genes.

Acquisition of immune‑related risk lncRNAs and PCA

The “limma” R software package was used to screen differential genes between lncRNA and the eight genes obtained, and the absolute value of cor > 0.4. When cor is positive, it is related to positive regulation and vice versa. \( P < 0.001 \) indicates statistical significance.

We then conducted univariate and multivariate analyses on the survival prognostic factors of these immune-related lncRNAs using the “survival” R package. In univariate Cox analysis, HR (95% CI) > 1 indicated high risk, and HR < 1 indicated low risk. A \( P \) value (survival time) < 0.0002 indicated statistical significance.

Correlation between the prognosis model and clinicopathological characteristics

Using the “pheatmap” package for analysis, the median risk score of the entire discovery group was set to 1.0 as the cutoff value, and 343 patients with HCC were divided into low- and high-risk groups. The relationship between the expression values of the five lncRNAs and the risk score, survival time, and status of each patient was analyzed. The “survival” and “survminer” packages were used to obtain Kaplan–Meier curves and \( P < 0.01 \) was considered statistically significant. The R software package “survival ROC” was used to study the prognostic value of gene signatures over time (Lambert and Chevret 2016). The R software package “ggpubr” was used to analyze the relationship between risk gene expression and clinical characteristics (such as T, G, and S).

Building and validating a predictive nomogram

Nomogram (Iasonos et al. 2008) was established by including all the prognostic factors. A calibration chart and the consistency index were used to study the calibration and identification of the nomogram (using the bootstrap method with 1000 replicates). Continuous and categorical variables were also considered. The maxscale parameter specified the maximum score. The prediction probability was converted to linearity. More than two survival probability axes were plotted. The list function was wired. An axis scale was set for survival or risk. Based on the stepwise Cox regression model, prognostic nomograms predicting the total survival times of 1, 2, and 3 years were established.

Gene set enrichment analysis

GSEA analysis was conducted according to sample grouping and gene expression profiles. In this mode, two files need to be input. The first file is the sample expression matrix, and the second file is the sample grouping file. Grouping files were arranged according to group, and enrichment analysis between the two groups was supported by default. GSEA analysis was performed after grouping according to the expression level of a gene in the expression profile. GSEA analysis was conducted after ranking according to the expression levels of specific genes. The input file consisted of two columns: the first column was the gene name and the second column was the gene expression level. Output the result. GSEA was supported by the Broad Institute website (https://www.gsea-msigdb.org/gsea/index.jsp) (Croken et al. 2014), which includes versions compatible with Java and R. All GSEA analyses were performed using Java GSEA software.

Cell culture

HepG2, SMMC-7721, Huh-7, and HL7702 cells were purchased from Changsha Abiowell Biology. The cells were digested with 0.25% trypsin digestion solution (C0201, Beyotime), and cell suspensions were prepared in RPMI1640 complete medium (04–001-1ACS, Gibco). HepG2 cells were transfected with the NC null plasmid and the PRRT3-AS1 inhibitor plasmid. All plasmids were purchased from Honorgene.

Quantitative real‑time PCR

Total RNA from cells was extracted using TRIzol (15,596,026, Thermo Fisher Scientific, Waltham, MA, USA). The sample RNA was reverse transcribed to cDNA according to the instructions of the reverse transcription kit (cw2569, Kangwei Century Company, China). The reaction conditions were denaturation at 95 °C for 10 min, denaturation at 94 °C for 15 s, and annealing at 60 °C for 30 s, for a total of 40 cycles. β-actin was used as the internal reference gene. The primer internal reference was β-actin. The primer
sequences are shown in Table 1. With 2 μg cDNA as a template, the relative quantitative method (2 − ΔΔCt method) was used to calculate the relative transcription level of the target gene: ΔΔCt = Δ experimental group - Δ Control group, ΔCt = Ct (target gene)-Ct (β-actin).

**Western blot**

The protein supernatant was mixed with the loading buffer, boiled with water for 5 min, and put into an icebox for later use. Separation gels with concentrations of 10% and 12% and concentrated gels with concentrations of 4.8% were prepared. After the gel was set up, 10–20 μl of protein solution was added to each well. Electrophoresis was started at a constant voltage of 75 V for 80 min. The protein was transferred to a nitrocellulose carbon membrane by applying a voltage. The membrane was then sealed with 5% skim milk powder for 90 min. After closure, the membrane was incubated with primary antibodies against CD24 (10,600–1-AP, Proteintech), THY1 (66,766–1-Ig, 1:5000, Proteintech), LYN (ab32398, 1:1000, Abcam), CD47 (ab218810, 1:1000, Abcam), TRAF2 (ab126758, 1:1000, Abcam), and β-actin (66,009–1-Ig, 1:5000, Proteintech) for 90 min. The membrane is then mixed with HRP Goat Anti-Mouse IgG (SA00001-1, 1:5000, Proteintech) or HRP Goat Anti-Rabbit IgG (SA00001-2, 1:6000, Proteintech) and incubated at room temperature for 90 min. The strips on the membrane were visualized using SuperECL Plus ultra-sensitive luminescence solution (K-12045-D50, Advansta) and X film. β-actin was used as an internal reference to analyze the expression level of the target protein.

**Enzyme-linked immunosorbent assay**

IL-1β (40,000,910, KE00021), IL-6 (40,001,125, KE00139), IL-10 (40,001,097, KE00170), TNF-α (40,001,179, KE00154), and TGF-β1 (40,001,153, KE00002) were purchased from Proteintech. The ELISA kits were purchased from Wuhan Huamei Biological Engineering Co., LTD. We added 50 μl of termination solution to each well sequentially. Within 5 min after the termination of the reaction, the optical density (OD) of each well was measured sequentially at a wavelength of 450 nm using an enzyme labeling instrument.

**Counting kit**

Cell proliferation was measured using a CCK-8 kit (AWC0114a, abiowell, China). Cells in the logarithmic growth phase were inoculated into 96-well plates at 5 × 10^3 cells/well and cultured for 3 days. Each group was set up with three multiple wells for corresponding intervention treatment according to the experimental groups and cultured for 24 h. After the corresponding incubation time, the medium was absorbed and replaced with CCK8 working solution (100 μl of medium and 10 μl of CCK8 initial solution), and incubated for 4 h in a 5% CO2 incubator at 37°C. Absorbance at 450 nm was measured using a microplate reader.

**Flow cytometry**

The cells were centrifuged at 1500 rpm for 5 min and washed once with PBS. The cells were suspended in 500 μl of Binding Buffer. Five microliters Annexin V-FITC (KGA108, KeyGEN, China) was added and mixed. Propidium Iodide (5 μl) was added and mixed. The reaction time was 5–15 min at room temperature in the dark. Finally, it was detected using a flow cytometer (a00-1–1102, Beckman, America).

**Transwell**

A total of 500μL 10% FBS complete medium was placed in the lower layer of the chamber. The pore size is 8 μm. The intervention cells were digested with 0.25% trypsin digestion solution (C0201, Beyotime, China), a serum-free basal medium was used to prepare the cell suspension, and the cell density was adjusted to approximately 1 × 10^5/ml. A total of 100 μl of each sample was absorbed and added to the upper chamber, and a complete medium containing 10% FBS (04–001-1ACS, Gibco, USA) was added to the lower chamber for 48 h.

**Statistical analysis**

R (version 3.6.2) was used for data analysis. The “limma” package was used to identify the DEGs. The threshold value of DEGs was set to log2 fold change (FC) > 1.0 and P value < 0.05. The “survival” and “survminer” R packages were used to explore the best-dividing line of risk score, and draw the Kaplan–Meier survival curve (P < 0.001). The “CorrPlot” R package was used to construct the correlation matrix between different genes. In univariate Cox analysis, HR (95% CI) > 1 indicated high risk, and HR < 1 indicated low risk. P value (survival time) < 0.0002 indicated statistical significance. The R software package “survival ROC” was used to study the prognostic value of gene signatures over time. The R software package “ggpubr” was used to

| Table 1 Primer sequence |
|-------------------------|
| Gene                  | Sequences (5'–3') |
| PRRT3-AS1              | F: GAAATGTATTAGCCCCGGTGC  
|                        | R: GTCTCACCTTCAGAGGGGGCG |
| β-actin                | F: ACCCTGAAAGTACCCCATCGAG  
|                        | R: AGCACAGGCTGGATAGCAAC   |
analyze the relationship between risk gene expression and clinical characteristics. Unpaired Student’s t-test was used to compare the data of two groups that did not have a one-to-one correspondence. One-way ANOVA and Tukey’s post hoc tests were used to compare the data among the three groups. Differences were considered statistically significant at P < 0.05.

Results

Differentially expressed immune genes co-expressed in TCGA and GEO databases and their interactions

After raw data pre-processing, samples from 343 patients with HCC (with a survival time longer than 30 days) and 50 healthy subjects from TCGA (using only patient samples) were used. An additional 81 HCC samples were retrieved from the GSE27150 data set. These samples were analyzed in the immune-related database via the gene set enrichment analysis (GSEA; https://www.gsea-msigdb.org/gsea/index.jsp). We used two datasets, IMMUNE_RESPONSE (M13664) and IMMUNE_SYSTEM_PROCESS (M19817), to obtain immune-related genes (IRGs). A total of 331 IRGs were downloaded from the TCGA database. Then, by selecting the intersecting genes, 42 IRGs co-expressed in TCGA and GEO databases were obtained. Subsequently, a total of 24 differentially expressed genes (DEGs) were identified, including CD24, THY1, LYN, CD47, and TRAF2.

The expression of each gene is displayed on a heat map (Fig. 1A). The volcano map shows the distribution of genes, including 19 upregulated and 5 downregulated genes in HCC (Fig. 1B).

Subsequently, a protein–protein interaction (PPI) network of the DEGs was constructed and visualized using the STRING (Search Tool for the Retrieval of Interacting Gene) database, as shown in (Fig. 1C). A total of 22 nodes and 47 edges were present in the PPI network, and the histogram shows the node number for each gene (Fig. 1D). We discovered central genes such as IL10, CD28, CCL2, CXCL12, CD24, CCR1, IL15, THY1, and TNFRSF14. Therefore, utilizing the TCGA and GEO databases, we found 24 IRGs that were differentially expressed in HCC and healthy subjects. Among these, IL10, CD28, CCL2, CXCL12, CD24, CCR1, IL15, THY1, and TNFRSF14 are central genes.

The acquisition of risk genes and their correlation with clinical prognosis and previous correlation

In order to obtain IRGs related to the clinical prognosis from the TCGA database, univariate and lasso Cox regression analyses were performed. P < 0.001 was considered statistically significant in univariate Cox regression analysis. The prognostic gene signature was calculated as follows: risk score = (0.0156018677981362 × TRAF2 expression) + (− 0.021929917526684 × MIA3 expression) + (0.00172601388297704 × CD24 expression) + (0.0193704943364634 × LYN expression) + (0.11198192547848 × MAP3K7 expression) + (0.01244164026282615 × CCR1 expression) + (0.0235215834428141 × CKLF expression) + (0.101112508330369 × HELLS expression). The R package “survival” was used to explore the best-dividing lines of patient risk score, and the Kaplan–Meier survival curves were constructed. Eight risk genes (TRAF2, MIA3, CD24, LYN, MAP3K7, CCR1, CKLF, and HELLS) were obtained (Table 2). Besides, the Kaplan–Meier survival curve demonstrated that these eight genes were related to prognosis in both TCGA and GEO databases and showed that the high-risk group had a lower 5-year survival rate than the low-risk group of patients with HCC (P < 0.001) (Fig. 2A, B). Next, construction and analysis of the correlation matrix between different genes revealed that IRG CCR1 had the highest correlation with the LYN (tyrosine kinases) gene, and Spearman’s rank correlation coefficient (rs) was calculated to be 0.61. All other rs values were ≤ 0.5 (Fig. 2C). The “CorrPlot” R package was utilized to create the figure. Therefore, we found 8 genes related to prognosis and the correlations between them. These 8 genes were TRAF2, MIA3, CD24, LYN, MAP3K7, CCR1, CKLF, and HELLS.

Acquisition of immune-related risk IncRNAs and principal component analysis

To identify immune-related IncRNAs, the IncRNAs in the TCGA database were co-expressed with eight previously related IRGs and checked for correlation using the “limma” R package. P values were considered statistically significant when the total score was < 0.001. Therefore, we obtained 678 differentially expressed IncRNAs, 11 of which were related to negative regulation, and the rest were related to positive regulation.

Subsequently, univariate and multivariate analyses of survival prognostic factors for these immune-related IncRNAs were performed using the “survival” R package. In univariate Cox analysis, hazard ratio (HR; 95% confidence interval (CI)) > 1 indicated high risk, and HR < 1 indicated low risk. A P value (survival time) of < 0.0002 was considered statistically significant. A total of 14 related IncRNAs were obtained, of which TMEM220-AS1 was at low risk, and the other 13 were at high risk (Fig. 3A). Multivariate Cox analysis was performed to obtain five immune-related...
Then, we conducted principal component analysis (PCA) of immune genes, immune-related lncRNAs, and immune-related risk lncRNAs to test their ability to identify the prognosis of patients with HCC. The results show that immune genes (green and red circles) were clustered together, and immune-related lncRNAs (red-green and red circles) were more dispersed. Immune-related risk lncRNAs (red-green and red circles) were essentially separated (Fig. 3B). Therefore, we found five immune-related lncRNAs (AC099850.3, AL031985.3, PRRT3-AS1, AC023157.3, and MSC-AS1) in TCGA and GEO databases.

### Correlation between prognosis model and clinicopathological characteristics

The TCGA database included 343 patients with information on their clinical features, such as gender, age, T (tumor), grade, clinical stage, lymph node metastasis, and survival time and status. Selecting 1.0 to be the
median risk score for the entire discovery group was used as the cutoff value, and 343 patients were classified as high- or low-risk patients (Fig. 4B). Figure 4C shows the survival status of HCC patients in different groups. Figure 4A shows the expression levels of the five lncRNAs (MSC-AS1, AC099850.3, PRRT3-AS1, AL031985.3, and AC023157.3) of each patient. Strictly, we can know that the expressions of MSC-AS1, AC099850.3, PRRT3-AS1, AL031985.3, and AC023157.3 in high-risk patients are higher than those in low-risk patients.

**Fig. 2** Kaplan–Meier curves and interaction with eight genes: Kaplan–Meier curves were used to analyze the overall survival rate of patients with hepatocellular carcinoma in high-risk and low-risk groups from, TCGA (The Cancer Genome Atlas; A) and GEO (Gene Expression Omnibus; B). C Terpene correlation plot. Visual representation of the correlation matrix; darker blue and larger dot size indicate a stronger positive correlation, whereas darker red indicates a stronger negative correlation

**Patterns for patients in high- and low-risk groups were determined using the five-immune-related IncRNA signature**

Further analysis of the data from the Kaplan–Meier curves revealed that the survival probability of high-risk patients was significantly lower than that of low-risk patients, and the \( P \) value of the log-rank test was \( 3.033 \times 10^{-5} \) (Fig. 5A). In addition, the area under the curve (AUC) of patient survival time from the TCGA database revealed that the AUC of risk genes was 0.733, the AUC of the stage was 0.743,
**Fig. 3** Acquisition of immune-related risk long non-coding RNAs (lncRNAs) and principal component analysis (PCA). **A** Forest map of hazard ratio (HRs) of immune-related lncRNAs. **B** PCA for all immune genes, immune lncRNAs, and risk lncRNAs; red dots represent high-risk genes, and green dots represent low-risk genes.

**Table 3** Five immune-related lncRNAs selected as prognosis-associated factors in hepatocellular carcinoma

| Gene symbol | Chromosome             | P value | Hazard ratio | Coefficient |
|-------------|------------------------|---------|--------------|-------------|
| AC004656.1  | Chr 17:59,197,566–59,215,226 | <0.001  | 1.691(1.286–2.223) |             |
| SNHG3       | Chr 17:59,197,566–59,215,226 | <0.001  | 1.526(1.260–1.847) |             |
| AC027097.1  | Chr 17:59,197,566–59,215,226 | <0.001  | 3.639(2.135–6.204) |             |
| LUCAT1      | Chr 17:59,197,566–59,215,226 | <0.001  | 1.702(1.354–2.140) |             |
| PIK3CD–AS2  | Chr 17:59,197,566–59,215,226 | <0.001  | 1.632(1.299–2.052) |             |
| AL117338.3  | Chr 17:59,197,566–59,215,226 | <0.001  | 2.275(1.538–3.364) |             |
| AC099985.3  | Chr 17:59,197,566–59,215,226 | <0.001  | 1.590(1.323–1.911) |             |
| NRAV        | Chr 17:59,197,566–59,215,226 | <0.001  | 2.166(1.618–2.899) |             |
| AL031985.3  | Chr 17:59,197,566–59,215,226 | <0.001  | 3.461(2.338–5.123) |             |
| PRRT3–AS1   | Chr 17:59,197,566–59,215,226 | <0.001  | 1.439(1.190–1.740) |             |
| TMEM220–AS1 | Chr 17:59,197,566–59,215,226 | <0.001  | 0.530(0.397–0.708)  |             |
| AC023157.3  | Chr 17:59,197,566–59,215,226 | <0.001  | 1.906(1.368–2.655) |             |
| MSC–AS1     | Chr 17:59,197,566–59,215,226 | <0.001  | 1.726(1.383–2.153) |             |
| LINC01138   | Chr 17:59,197,566–59,215,226 | <0.001  | 2.140(1.509–3.036) |             |
and the AUC of tumor (T) stage was 0.752, indicating that immune-related lncRNA signatures have a moderate potential to monitor survival (Fig. 5B). Compared to other clinicopathological features, our model achieved the largest AUC value, which also reflects its excellent predictability.

We then analyzed the relationship between the expression of risk genes and clinical features such as T, grade (G), and stage (S). In the TCGA database, the expression of AC023157.3, AC099850.3, and AL031985.3 was found to be associated with T, AC099850.3, AL031985.3, and PRRT3-AS1 were found to be associated with grade, and AC023157.3 and AC099850.3 were found to be related to the stage (P < 0.05; Fig. 5C). Therefore, we verified that the lncRNAs AC023157.3, AC099850.3, AL031985.3, and PRRT3-AS1 were related to clinical features such as T, G, and S.

We then analyzed the relationship between the expression of risk genes and clinical features such as T, grade (G), and stage (S). In the TCGA database, the expression of AC023157.3, AC099850.3, and AL031985.3 were found to be associated with T, AC099850.3, AL031985.3, and PRRT3-AS1 were found to be associated with grade, and AC023157.3 and AC099850.3 were found to be related to the stage (P < 0.05; Fig. 5C). Therefore, we verified that the lncRNAs AC023157.3, AC099850.3, AL031985.3, and PRRT3-AS1 were related to clinical features such as T, G, and S.
Building and validating the predictive nomogram

TCGA dataset included 343 patients with complete clinical information. Based on the stepwise Cox regression model, a prognostic nomogram that predicts the overall survival rate at 1, 2, and 3 years was established (Fig. 6). Parameters such as the risk score, age, sex, grade, stage, and tumor node metastasis (TNM) were included in the nomogram. The total nomogram score showed the predicted survival of patients with HCC at 1, 2, and 3 years. The 1-, 2-, and 3-year survival rates were closely related to the risk score, which could help guide clinical decision-making.

Gene set enrichment analysis

We further analyzed immune-related lncRNAs and enrichment scores from the GSEA analysis. This analysis used permutation lists and leading-edge proteins were considered hits before the enrichment score for the gene set. GSEA was performed to elucidate the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway and Gene Ontology (GO). KEGG pathway analysis showed that in the high-risk group, drug metabolism cytochrome P450 (CYPs), and metabolism of xenobiotics by CYPs, mTOR, and p53 pathways were found to be enriched. In low-risk patients, the cell cycle and fatty acid metabolism pathways were found to be enriched (Fig. 7A). For GO pathways, most enriched pathways in high-risk patients were linked to gene signaling, high-density lipoprotein, monocarboxylic acid catabolic, organic acid catabolic, and RNA splicing. In addition, in low-risk patients, the pathways were enriched in benzene-containing compound metabolism, histone deacetylase complex, and mRNA binding (Fig. 7B).
Effects of PRRT3-AS1 on the inflammatory response of HepG2 cells

From the above-screened lncRNAs, PRRT3-AS1 was selected for further functional studies. First, the hepatocellular carcinoma cell lines were screened. qRT-PCR data showed that PRRT3-AS1 was elevated in human hepatocellular carcinoma cells SMMC-7721, HepG2, and Huh-7 compared to the normal hepatocellular line HL7702, with the most significant elevation of PRRT3-AS1 in HepG2 cells (Fig. 8A). Therefore, HepG2 cells were selected for subsequent experiments. The results of the data showed that PRRT3-AS1 expression was reduced in HepG2 cells transfected with PRRT3-AS1 inhibitory plasmid, which indicated that the cell model was successfully constructed (Fig. 8B). To further understand the effect of PRRT3-AS1 on the inflammatory response in HepG2 cells. Compared with the control and si-NC groups, the levels of anti-inflammatory factors TGF-β1 and IL-10 were increased and the levels of pro-inflammatory factors IL-1β, TNF-α, and IL-6 were decreased in the si-PRRT3-AS1 group (Fig. 8C). This indicated that inhibiting the expression of PRRT3-AS1 in HepG2 cells could reduce the inflammatory response of the cells.

The effect of PRRT3-AS1 on HepG2 cell function may be through mRNAs

The above results indicated that PRRT3-AS1 could inhibit the inflammatory response of HepG2 cells. For mRNAs deleted from the abovementioned GEO and TCGA databases, the protein expression of CD24, THY1, LYN, CD47, and TRAF2 was suppressed by the knockdown of PRRT3-AS1 (Fig. 9A). The impaired function of HepG2 cells may be regulated by PRRT3-AS1, which affects the expression of these mRNAs. The results of the CCK-8 assays showed that compared with the control and si-NC groups, HepG2 cells in the si-PRRT3-AS1 group showed significantly decreased proliferation ability (Fig. 9B), the number of migrating and invasive cells (Fig. 9C and D), and the apoptosis rate increased significantly (Fig. 9E). In summary, CD24, THY1, LYN, CD47, and TRAF2 could be targets of PRRT3-AS1 regulation in HepG2 cells.
Discussion

As HCC is a heterogeneous disease with a poor prognosis and high mortality, it is crucial to investigate novel biomarkers that can effectively predict the survival of patients with HCC. In recent years, with the development of sequencing, many human transcripts without protein-coding potential called non-coding RNAs (ncRNA), including microRNA (miRNA), small nucleolar RNA (snoRNA), long non-coding RNA (lncRNA), and circular RNA (circRNA) have been identified. LncRNAs are known to be more abundant than standard protein-coding mRNA (Shi et al. 2013; Chen and Conn 2017). They are known to play a role in gene expression, gene imprinting, chromatin remodeling, transcription, post-transcription, and epigenetic regulation (Momen-Heravi and Bala 2018; Zhang et al. 2019b; Zheng et al. 2022).

A variety of lncRNAs are abnormally expressed in HCC and participate in tumor metastasis, recurrence, and prognosis. These include interactions with DNA, RNA, or proteins to form a complex that regulates target gene expression through various mechanisms (Liu et al. 2019a; Lou et al. 2019; Shi et al. 2019; Wang et al. 2019; Wei et al. 2019; Zhang et al. 2019a). LncRNAs also have important roles in cancer immunity, such as antigen release, immune cell migration and filtration, antigen presentation, and immune activation (Atianand et al. 2017; Chen et al. 2017). Immunotherapy for liver cancer has also received extensive attention (Sim and Knox 2018; Dal Bo et al. 2020). TCGA is an open database containing samples from patients with various malignancies. The GEO database has been widely used in the bioinformatic analysis because it is a comprehensive database that stores large amounts of gene expression data. In this study,
Fig. 8 PRRT3-AS1 can inhibit the inflammatory response of HepG2 cells.  

A qRT-PCR was used to measure PRRT3-AS1 expression in HL7702, SMMC-7721, HepG2, and Huh-7 cells.  

B qRT-PCR was used to detect the expression levels of PRRT3-AS1.  

C ELISA was used to determine the levels of TNF-α, IL-6, TGFβ1, IL-10, and IL-1β.  

*P < 0.05, compared with the HL7702 group; #P < 0.05, compared with the control group.  

Comparisons among multiple groups were analyzed using one-way ANOVA, followed by Tukey’s post hoc test. n = 3

we downloaded RNA sequencing data from TCGA and GEO databases and obtained the lncRNA expression profiles of patients with HCC in the dataset for analysis.

Liu et al. developed prognostic and diagnostic classification tool for HCC. Analyzing the data in the database through the prognostic model related to the predicted overall survival of HCC may be helpful for clinical decision-making in individual treatments (Liu et al. 2019b). In this study, we used a TCGA training cohort including 343 patients with HCC, examined the relationship between the expression of immune-related lncRNAs and the prognosis of patients with HCC, and verified their effectiveness using the GSE27150 dataset. First, by co-expressing genes from the two databases and performing Cox regression analysis, 24 immunity-related genes were obtained. Univariate and lasso Cox regression analyses were used to identify related immune genes to predict and construct prognostic gene signatures. Eight immunity-related genes (TRAF2, MIA3, CD24, LYN,
MAP3K7, CCR1, CKLF, and HELLS) were identified as prognostic factors for HCC. Among them, TRAF2 has been proven to promote the development of liver cancer (Schneider et al. 2017). CD24 was associated with a poor prognosis in HCC patients (Maimaitiming et al. 2020). MAP3K7 was a potential marker of HCC (Cheng et al. 2019). Earlier studies have shown that CCR1 was abnormally expressed in HCC (Yang et al. 2006). In addition, CKLF1 was overexpressed in HCC in vitro and in vivo experiments (Liu et al. 2019c). Similarly, HELLS has also been proven to be a marker of HCC (Law et al. 2019). However, the research on MIA3 and LYN is temporarily insufficient in HCC. Then, through co-expression, univariate, and multivariate analyses, five immune-related lncRNAs (AC099850.3, AL031985.3, PRRT3-AS1, AC023157.3, and MSC-AS1) affecting the prognosis were obtained. Previous bioinformatics analysis results have indicated that these five lncRNAs were related to the prognosis of HCC (Jia et al. 2020; Kong et al. 2020; Wu et al. 2020; Xu et al. 2021a, b). At the same time, Cao et al. verified that MSC-AS1 was related to the development of HCC in vivo and vitro (Cao et al. 2020). Drug response analysis from Zhao Z revealed that PRRT3-AS1 is a potential resistance biomarker for paclitaxel BRCA treatment in breast carcinomas (Zhao, 2021 #1394). Furthermore, PRRT3-AS1 and AL031985.3 were identified as immune-related prognostic lncRNAs in HCC patients according to study from Liang R (2019 #1395). Zhang P found that Five lncRNAs including PRRT3-AS1 were found related to focal adhesion, extracellular matrix receptor interaction, and mitogen-activated protein kinase signaling pathways in prostate cancer. Fan L also revealed that silencing of lncRNA PRRT3-AS1 can upregulate apoptosis and autophagy downregulate migration and invasion of PC cells through the mTOR signaling pathway. Our results unveiled that the overall survival rate of patients was found to be significantly correlated with these lncRNAs. Through multiple Cox regression analysis and risk-scoring methods, five immune-related lncRNA signals were obtained, and patients with HCC were divided into high-risk and low-risk groups. Survival analysis showed that the overall survival rates of the two groups were significantly different. Since there was a possibility of overturning or five lncRNA signatures being false positive, we used a separate cohort of 343 patients with HCC to further validate the prognostic value of the five lncRNA signatures. The results of the independent verification showed that the five lncRNA signatures had good repeatability and robustness in predicting the prognosis of patients with HCC. However, due to molecular heterogeneity, traditional clinical predictors had limited success in predicting the survival of patients with HCC. The results of multivariate analysis indicate that the five lncRNA signatures did not depend on traditional clinical risk factors and molecular characteristics. Among them, PRRT3-AS1 can affect the proliferation, migration, invasion, and apoptosis of HepG2 cells. Silencing PRRT3-AS1 can inhibit the protein expression of CD24, THY1, LYN, CD47, and TRAF2. The role of PRRT3-AS1 and its downstream target genes in HCC deserves further study in the future. The mechanism of PRRT3-AS1 was not explored at the clinical level, which is a direction for further research in the future. Other genes not validated such as AC099850.3, AL031985.3, AC023157.3, and MSC-AS1 deserve to be investigated in more depth as the technology evolves.

Finally, the results of GSEA analysis indicated that in the high-risk group, drug metabolism of CYPs, and metabolism of xenobiotics by CYPs, mTOR, and p53 pathways were found to be enriched in the KEGG pathway. Studies have shown that HCC patients with downregulated CYPs might be more susceptible to drug toxicity, and the late stage of HCC was related to severe changes in drug metabolism (Nekvindova et al. 2020). In addition, the p53 and mTOR signaling pathways are closely related to tumors, such as HCC and NSCLC (Kong et al. 2019; Rebouissou and Nault 2020). This reflected the consistency between our research and previous studies. In terms of GO pathways, the high-density lipoprotein, monocarboxylic acid catabolic, and organic acid catabolic pathways were enriched in high-risk patients. Increasing evidence has verified that changes in lipid metabolism are the hallmarks of cancer (Tian et al. 2019). Various monocarboxylic acids promote HCC progression through various signaling pathways (Fan et al. 2018). Therefore, the progress of HCC through high-density lipoprotein, monocarboxylic acid catabolic, and organic acid catabolic remained to be discovered.

**Conclusion**

In summary, this study identified five immune-related lncRNA signals that can predict the prognosis of patients with HCC. Our findings also provide a basis for the selection of therapeutic targets and prognostic molecular markers.

**Supplementary Information** The online version contains supplementary material available at https://doi.org/10.1007/s10142-023-01019-x.
Acknowledgements Thank for contributors of TCGA and GEO for sharing the hepatocellular carcinoma data.

Author contribution AM conceived, designed, and investigated the research work; LX and HD supervised data collection and conducted data analysis; AM, LX, NST, and HD prepared the original draft; HZ reviewed and edited the manuscript. All authors contributed to editorial changes in the manuscript. All authors read and approved the final manuscript.

Data availability The GSE27150 gene expression profiles came from GEO (https://www.ncbi.nlm.nih.gov/geo). The statistical analyses used the R software (https://www.r-project.org/). The immune-related gene (IRGs) data came from the ImmPort database (https://immport.niaid.nih.gov). The TCGA data came from the TCGA database (https://portal.gdc.cancer.gov/). The immune-related genes (IRGs) data came from the ImmPort database (https://immport.niaid.nih.gov). The statistical analyses used the R software (https://www.r-project.org/).

Declarations

Competing interests The authors declare no competing interests.

References

Atianand MK, Caffrey DR, Fitzgerald KA (2017) Immunobiology of long noncoding RNAs. Annu Rev Immunol 35:177–198. https://doi.org/10.1146/annurev-immunol-041015-035459

Bray F, Ferlay J, Soerjomataram I et al (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68:394–424. https://doi.org/10.3322/caac.21492

Cao C, Zhong Q, Lu L et al (2020) Long noncoding RNA MSC-AS1 promotes hepatocellular carcinoma oncogenesis via inducing the expression of phosphoglycerate kinase 1. Cancer Med 9:5174–5184. https://doi.org/10.1002/cam4.3080

Chen JA, Conn S (2017) Canonical mRNA is the exception, rather than the rule. Genome Biol 18:133. https://doi.org/10.1186/s13059-017-1268-1

Chen YG, Satpathy AT, Chang HY (2017) Gene regulation in the microsystem by long noncoding RNAs. Nat Immunol 18:962–972. https://doi.org/10.1038/ni.3771

Chen Z, Chen X, Lei T et al (2020) Integrative analysis of NSCLC identifies LINCO1234 as an oncogenic IncRNA that interacts with HNRNPA2B1 and regulates miR-106b biogenesis. Mol Ther 28:1479–1493. https://doi.org/10.1016/j.ymthe.2020.03.010

Cheng JS, Tsai WL, Liu PF et al (2019) The MAP3K7-mTOR axis promotes tumourigenesis by inhibiting ANXA2 phosphorylation in hepatocellular carcinoma. Front Oncol 9:474. https://doi.org/10.3389/fonc.2019.00474

Croken MM, Qiu W, White MW et al (2014) Gene Set Enrichment Analysis (GSEA) of Toxoplasma gondii expression datasets links cell cycle progression and the Bradyzoite developmental program. BMC Genomics 15:515. https://doi.org/10.1186/1471-2164-15-515

Dal Bo M, De Mattia E, Baboci L et al (2020) New insights into the pharmacological, immunological, and CART-cell approaches in the treatment of hepatocellular carcinoma. Drug Resist Update 51:100702. https://doi.org/10.1016/j.drup.2020.100702

Fan Q, Yang L, Zhang X et al (2018) Autophagy promotes metastasis and glycolysis by upregulating MCT1 expression and Wnt/β-catenin signaling pathway activation in hepatocellular carcinoma cells. J Exp Clin Cancer Res 37:9. https://doi.org/10.1186/s13046-018-0657-y

Forner A, Reig M, Bruix J (2018) Hepatocellular carcinoma. Lancet 391:1301–1314. https://doi.org/10.1016/s0140-6736(18)30010-2

Franceschini A, Szklarczyk D, Frankish A et al (2013) STRING v9.1: protein interaction networks, with increased coverage and integration. Nucleic Acids Res 41:D808-815. https://doi.org/10.1093/nar/gks1094

Gao F, Cai Y, Kapranov P et al (2020) Reverse-genetics studies of lncRNAs-what we have learnt and paths forward. Genome Biol 21:93. https://doi.org/10.1186/s13059-020-01994-5

Iasonos A, Schrag D, Raj GV et al (2008) How to build and interpret a nomogram for cancer prognosis. J Clin Oncol 26:1364–1370. https://doi.org/10.1200/JCO.2007.12.9791

Jia Y, Chen Y, Liu J (2020) Prognosis-predictive signature and nomogram based on autophagy-related long non-coding RNAs for hepatocellular carcinoma. Front Genet. 11:608668. https://doi.org/10.3389/fgene.2020.608668

Kong N, Tao W, Ling X et al (2019) Synthetic mRNA nanoparticle-mediated restoration of p53 tumor suppressor sensitizes p53-deficient cancers to mTOR inhibition. Sci Transl Med 11. https://doi.org/10.1126/scitranslmed.aaw1565

Kong W, Wang X, Zuo X et al (2020) Development and validation of an immune-related lncRNA signature for predicting the prognosis of hepatocellular carcinoma. Front Genet 11:1037. https://doi.org/10.3389/fgene.2020.01037

Lambert J, Chevret S (2016) Summary measure of discrimination in survival models based on cumulative/dynamic time-dependent ROC curves. Stat Methods Med Res 25:2088–2102. https://doi.org/10.1177/096228021551571

Law CT, Wei L, Tsang FH et al (2019) HELLS regulates chromatin remodeling and epigenetic silencing of multiple tumor suppressor genes in human hepatocellular carcinoma. Hepatology 69:2013–2030. https://doi.org/10.1002/hep.30414

Li C, Jiang W, Xu Y (2017) Omics and bioinformatics: time for new data analysis approaches? OMICS 21:749. https://doi.org/10.1089/omi.2017.0120

Li X, Xiao X, Chang R et al (2020) Comprehensive bioinformatics analysis identifies IncRNA HCG22 as a migration inhibitor in esophageal squamous cell carcinoma. J Cell Biochem 121:468–481. https://doi.org/10.1002/jcb.29218

Li T, Zhou Y, Li D et al (2022) The role of genome-scale leukocyte long noncoding RNA in identifying acute aortic dissection. Signa Vitae 18:101–110. https://doi.org/10.22514/sv.2022.030

Liu D, Gao M, Wu K et al (2019a) LINCO0152 facilitates tumorigenesis in esophageal squamous cell carcinoma via miR-153–3p/FYN axis. Biomed Pharmacother 112:108654. https://doi.org/10.1016/j.biopha.2019.108654

Liu GM, Zeng HD, Zhang CY et al (2019b) Identification of a six-gene signature predicting overall survival for hepatocellular carcinoma. Cancer Cell Int 19:138. https://doi.org/10.1186/s12935-019-0858-2

Liu Y, Liu L, Zhou Y et al (2019c) CKLF1 enhances inflammation-mediated carcinogenesis and prevents doxorubicin-induced apoptosis via IL6/STAT3 signaling in HCC. Clin Cancer Res 25:4145–4154. https://doi.org/10.1158/1078-0432.Ccr-18-3510

Lou Y, Yu Y, Xu X et al (2019) Long non-coding RNA LUCAT1 promotes tumourigenesis by inhibiting ANXA2 phosphorylation in hepatocellular carcinoma. J Cell Mol Med 23:1873–1884. https://doi.org/10.1111/jcmm.14088

Maimaitiming A, Zhou X, Ma X et al (2020) Clinicopathological and prognostic value of plasma CD24 level in hepatocellular carcinoma. J Invest Surg 33:536–541. https://doi.org/10.1080/0962280213515509

Momen-Heravi F, Bala S (2018) Emerging role of non-coding RNA in oral cancer. Cell Signal 42:134–143. https://doi.org/10.1016/j.cellsig.2017.10.009

Nekvindova J, Mrkvicova A, Zubenova V et al (2020) Hepatocellular carcinoma: gene expression profiling and regulation of xenobiotic-metabolizing cytochromes P450. Biochem Pharmacol 177:113912. https://doi.org/10.1016/j.bcp.2020.113912

Ni W, Zhang Y, Zhan Z et al (2017) A novel lncRNA uc.134 represses hepatocellular carcinoma progression by inhibiting...
CUL4A-mediated ubiquitination of LATS1. J Hematol Oncol 10:91. https://doi.org/10.1186/s13045-017-0449-4
O’Connor HJ (2000) “Lasso” technique for retrieval of impacted biliary stent. Gastrointest Endosc 51:640–641
Rebouissou S, Nault JC (2020) Advances in molecular classification and precision oncology in hepatocellular carcinoma. J Hepatol 72:215–229. https://doi.org/10.1016/j.jhep.2019.08.017
Rheinheimer S, Heussel CP, Mayer P et al (2020) Oligoprogres¬sive non-small-cell lung cancer under treatment with PD-(L)1 inhibitors. Cancers (Basel) 12. https://doi.org/10.3390/cancers12041046
Rodriguez-Cerdeira C, Carnero Gregorio M, Lopez-Barcenas A et al (2017) Advances in immunotherapy for melanoma: a comprehensive review. Mediat Inflamm 2017:3264217. https://doi.org/10.1155/2017/3264217
Schneider AT, Gautheron J, Feoktistova M et al (2017) RIPK1 suppresses a TRAF2-dependent pathway to liver cancer. Cancer Cell 31:94–109. https://doi.org/10.1016/j.ccell.2016.11.009
Shi X, Sun M, Liu H et al (2013) Long non-coding RNAs: a new frontier in the study of human diseases. Cancer Lett 339:159–166. https://doi.org/10.1016/j.canlet.2013.06.013
Shi L, Hong X, Ba L et al (2019) Long non-coding RNA ZNFX1-AS1 promotes the tumor progression and metastasis of colorectal cancer by acting as a competing endogenous RNA of miR-144 to regulate EZH2 expression. Cell Death Dis 10:150. https://doi.org/10.1038/s41419-019-1332-8
Sim HW, Knox J (2018) Hepatocellular carcinoma in the era of immunotherapy. Curr Probl Cancer 42:40–48. https://doi.org/10.1016/j.curpproblcancer.2017.10.007
Thoma C (2018) Kidney cancer: combining targeted and immunotherapy. Nat Rev Urol 15:263. https://doi.org/10.1038/nrurol.2018.43
Tian Y, Yang B, Qiu W et al (2019) ER-residential Nogo-B accelerates liver cancer by inhibiting PTEN via autophagy cooperation to miR15a. Mol Cancer 17:94. https://doi.org/10.1186/s12943-018-0843-8
Xiong H, Ni Z, He J et al (2017) LncRNA HULC triggers autophagy via stabilizing Sirt1 and attenuates the chemosensitivity of HCC cells. Oncogene 36:3528–3540. https://doi.org/10.1038/onc.2016.521
Xu BH, Jiang JH, Luo T et al (2021a) Signature of prognostic epithelial-mesenchymal transition related long noncoding RNAs (ERLs) in hepatocellular carcinoma. Medicine (Baltimore) 100:e26762. https://doi.org/10.1097/md.00000000000026762
Xu Q, Wang Y, Huang W (2021b) Identification of immune-related lncRNA signature for predicting immune checkpoint blockade and prognosis in hepatocellular carcinoma. Int Immunopharmacol 92:107333. https://doi.org/10.1016/j.intimp.2020.107333
Yan C, Wei S, Han D et al (2020) LncRNA HULC shRNA inhibits miR-377–5p to suppress the growth and invasion of hepatocellular carcinoma in vitro and hepatocarcinogenesis in vivo. Ann Transl Med 8:1294. https://doi.org/10.21037/atm-20-5556
Yang X, Lu P, Fujii C et al (2006) Essential contribution of a chemokine, CCL3, and its receptor, CCR1, to hepatocellular carcinoma progression. Int J Cancer 118:1869–1876. https://doi.org/10.1002/ijc.21596
Zhang P, Dong Q, Zhu H et al (2019a) Long non-coding antisense RNA GAS6-AS1 supports gastric cancer progression via increasing GAS6 expression. Gene 696:1–9. https://doi.org/10.1016/j.gene.2018.12.079
Zhang Y, Li Z, Chen M et al (2019b) Identification of a new eight-long non-coding RNA molecular signature for breast cancer survival prediction. DNA Cell Biol 38:1529–1539. https://doi.org/10.1089/dna.2019.5059
Zheng S, Chen Z, Chen J et al (2022) A prognostic index model for predicting the prognosis of HCC through coexpression and ceRNA mechanism. Biomed Res Int 2020:3801748. https://doi.org/10.1155/2020/3801748
Xin X, Wu M, Meng Q et al (2018) Long noncoding RNA HULC accelerates liver cancer by inhibiting PTEN via autophagy cooperation to miR15a. Mol Cancer 17:94. https://doi.org/10.1186/s12943-018-0843-8
Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.