Four Immune-Related Long Non-coding RNAs for Prognosis Prediction in Patients With Hepatocellular Carcinoma

Muqi Li[1†], Minni Liang[2†], Tian Lan[3], Xiwen Wu[1], Wenxuan Xie[1], Tielong Wang[4,5,6], Zhitao Chen[4,5,6], Shunli Shen[1]* and Baogang Peng[1]*

[1] Department of Liver Surgery, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China, [2] Center of Surgery and Anaesthiology, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China, [3] Department of Pancreatobiliary Surgery, The First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China, [4] Organ Transplant Center, First Affiliated Hospital of Sun Yat-sen University, Guangzhou, China, [5] Guangdong Provincial Key Laboratory of Organ Donation and Transplant Immunology, Guangzhou, China, [6] Guangdong Provincial International Cooperation Base of Science and Technology (Organ Transplantation), Guangzhou, China

Background: Long non-coding RNA (LncRNA) plays an important role in the occurrence and development of hepatocellular carcinoma (HCC). This study aims to establish an immune-related LncRNA model for risk assessment and prognosis prediction in HCC patients.

Methods: Hepatocellular carcinoma patient samples with complete clinical data and corresponding whole transcriptome expression were obtained from the Cancer Genome Atlas (TCGA). Immune-related genes were acquired from the Gene Set Enrichment Analysis (GSEA) website and matched with LncRNA in the TCGA to get immune-related LncRNA. Least Absolute Shrinkage and Selection Operator (LASSO) regression was used for screening the candidate LncRNAs and calculating the risk coefficient to establish the prognosis model. Patients were divided into a high-risk group and a low-risk group depending on the median risk score. The reliability of the prediction was evaluated in the validation cohort and the whole cohort. GSEA and principal component analysis were used for function evaluation.

Results: A total of 319 samples met the screening criteria and were randomly distributed across the training cohort and the validation cohort. After comparison with the IMMUNE_RESPONSE gene set and the IMMUNE_SYSTEM_PROCESS gene set, a total of 3094 immune-related LncRNAs were screened. Ultimately, four immune-related LncRNAs were used to construct a formula using LASSO regression. According to the formula, the low-risk group showed a higher survival rate than the high-risk group in the validation cohort and the whole cohort. The receiver operating characteristic curves data demonstrated that the risk score was more specific than other traditional clinical characteristics in predicting the 5-year survival rate for HCC.

Conclusion: The four-immune-related-LncRNA model can be used for survival prediction in HCC and guide clinical therapy.

Keywords: hepatocellular carcinoma, long noncoding (lnc) RNAs, prognosis, survival, bioinformation analysis

INTRODUCTION

Hepatocellular carcinoma (HCC) is one of the most lethal cancers in the world. Although medical technology has improved in recent years, the 5-year survival rate and mortality rate of HCC have not significantly improved (Siegel et al., 2019). Surgery is the curative treatment for HCC, but less than 30% of patients have the opportunity of undergoing a radical operation upon diagnosis (European Association for the Study of the Liver et al., 2018). Therefore, the exploration of non-operative methods and an evaluation system for prognosis are urgently needed.

In addition to surgical procedures such as hepatectomy, radiofrequency ablation, and liver transplantation, non-operative therapies including targeted therapy with tyrosine kinase inhibitors and immunotherapy with programmed death 1 (PD-1) or programmed death 1 ligand (PD-L1) antibodies are now the mainstay treatments for advanced HCC. It has been proved that an immune response disorder in the cancer microenvironment plays an important role in the tumor development (Teng et al., 2015). Because of the general tolerance and immunosuppressive environment of the liver, once HCC has developed, this mechanism might protect the tumor and assist in its growth, further epithelial-mesenchymal transformation, and the development of a true tumor microenvironment, enhancing suppressive immunity such as the expression of PD-L1 and protecting cancer cells (Kubes and Jenne, 2018). Many cancers suppress T cell attacks by overexpressing inhibitory ligands to avoid suppression of the tumor by the immune system. Therefore, the low number of T cells in patients with HCC and their vulnerability to damage may be the reasons leading to the development of HCC (Jia et al., 2015). Over the past few years, great progress was made in the application of immune checkpoint inhibitors (ICI) such as cytotoxic T lymphocyte antigen 4 (CTLA-4), PD-1, and PD-L1 for the clinical treatment of HCC (Sprinzl and Galle, 2017). The CTLA-4 blocker tremelimumab was effective in approximately 17.6% of HCC patients with a hepatitis C virus infection, while the progression time was about 6.48 months (Sangro et al., 2013). In combination with local tumor ablation, tumor-associated antigens (TAAs) can be released from apoptotic or necrotic liver cancer tissue, which accelerates the activation of tumor-oriented T cells and immune synergism (Mizukoshi et al., 2011). On the other hand, the assertiveness of the PD-1 pathway can change the duration of T cell-APC or T cell-target cell contact and decrease the number of T-reg cells in the tumor to enhance the anti-tumor immune response (Fife et al., 2009; Francisco et al., 2009). Although PD1 antibody-based immunotherapy has made considerable strides in the treatment of HCC, there are still some deficiencies in the clinical application of these drugs. First, these drugs may cause complications such as diarrhea, rashes, lung disease, and even lethal complications (Zhong et al., 2016; Zhang X. et al., 2019). Second, their high price is beyond the reach of many families. Third, PD-1 antibody monotherapy appears to be effective only in a very small subset of patients, with no effect in the majority of patients (Shivaji et al., 2019). Hepatocellular carcinoma has

multiple immunophenotypes, which underscores the importance of further research on immune cell interaction in HCC and the development of individualized immunotherapy drugs (Hilmi et al., 2019; Llovet et al., 2019). The expression of PD-1 may be involved in the screening of sensitive populations with higher levels of PD-1/PD-L1 who have better histologic response rates and progression-free survival (PFS) and overall survival (OS) after immunotherapy (Patel and Kurzrock, 2015). Accurate screening criteria to maximize the benefits of immunotherapy are of great significance.

Long non-coding RNAs (LncRNAs) are a class of non-coding RNAs of more than 200 nucleotides in length (Ma et al., 2017). Long non-coding RNA plays a crucial part in gene expression, cell growth, differentiation, and other biological processes (Klec et al., 2019). At the same time, genome sequencing revealed that the expression of LncRNAs is different in different kinds of cancers (Jing et al., 2018; Nanni et al., 2020; Yn et al., 2020). In addition, studies indicated that LncRNAs may play a role in inhibiting or promoting the occurrence and development of cancers (He et al., 2014). Meanwhile, LncRNA may be a marker of tumor prognosis or an immunotherapy target (Li et al., 2020). Studies show that LncRNA plays a broad regulatory role in functional activities such as immune response, tumorigenesis, and tumor progression (Bhan et al., 2017; Yu et al., 2018). It was found that some of the immune-related LncRNAs regulate the biological behavior in diseases by regulating the immunosuppressive activity of bone marrow mesenchymal stem cells, the expression level of T/B cells and natural killer cells, etc. (Wang et al., 2015; Zheng et al., 2019; Ahmad et al., 2020). Long non-coding RNAs have been studied using high-throughput sequencing and bioinformatics methods, and their prognostic value has been evaluated by creating formulas for models (Wang et al., 2018; Yan et al., 2019; Shen et al., 2020). However, the relationship between immune-related LncRNAs and HCC prognosis has been studied less extensively.

In this study, we try to establish a prognosis model based on immune-related LncRNAs for predicting the prognosis and guiding the application of immunotherapy in HCC.

MATERIALS AND METHODS

Patient Cohort and Grouping

RNA sequencing data and corresponding clinical information of HCC patients were obtained from the Cancer Genome Atlas (TCGA)-Liver hepatocellular carcinoma (LIHC) of the TCGA database[1] (Gao et al., 2013; Li et al., 2015). All patients were diagnosed with HCC. Patients were excluded if their follow-up time was less than 30 days or their pathological stage was unclear. Overall, 319 patients met the screening criteria. The patients were randomly divided into a training cohort ($n = 160$) and a validation cohort ($n = 159$). A multi-LncRNA model of prognosis was established using the training cohort, and the

---

[1] https://cancergenome.nih.gov

validation cohort was used to test the predictive power of the equation. There is no difference in clinical characteristics between the training cohort and validation cohort.

The Expression of Genes and Relative Immune LncRNAs in the TCGA-LIHC Cohort

We integrated the human gene annotation file from the Gene Expression Omnibus (GEO) database and Ensemble website. Based on ensemble-ID, the genes obtained from the TCGA database were identified as protein-coding genes or non-coding genes. We then compared the contents with the data retrieved from the TCGA, enabling us to distinguish between LncRNA and other kinds of RNA from all gene expression (Zhang et al., 2012). The original data were processed by FPKM (fragments per kilobase million). This normalization method can avoid errors caused by differences in gene length and sequencing depth. The data acquired from the dataset were consolidated and processed by the "limma" package in R. In addition, 7009 LncRNAs were selected from the TCGA-LIHC project. A total of 331 immune-related genes were obtained by matching the gene sets from the TCGA with the IMMUNE_RESPONSE gene set (M19817) and the IMMUNE_SYSTEM_PROCESS gene set (M13664) in the molecular characterization database V7.0[2] (Subramanian et al., 2005; Wang et al., 2018; Sun et al., 2020). In total, 3094 immune-related LncRNAs were obtained through correlation analysis between the LncRNA expression and immune gene expression. According to the hazard ratio, the expression of LncRNA could be defined as positive or negative (cor > 0.4 and P value < 0.01). For patients with biological replicates, the final expression level was the mean expression of all repeated samples of LncRNA. The LncRNA was excluded from the expression matrix if it was missing in more than 20% of HCC patients, to avoid calculation bias (Yan et al., 2019).

Construction and Evaluation of an Immune-Related LncRNA Prediction Model

Univariate Cox regression analysis was used to determine the relationship between the level of immune-related LncRNA expression and OS. LncRNAs with a P value less than 0.01 are considered as significant prognostic LncRNAs. Least Absolute Shrinkage and Selection Operator (LASSO) Cox regression analysis and the "glmnet" package in R were used to analyze the best candidates and multiple immune LncRNA features for predicting the OS of patients (Zhang M. et al., 2019; Gupta et al., 2020). Based on this model, the LASSO regression coefficients weighted LncRNA expression levels were recruited to calculate the risk score for each patient.

The relationship between clinical features, risk score, and patient OS univariate was calculated by univariate independent prognostic analysis and multivariate independent prognostic analysis. The area under the curve (AUC) and the time-dependent receiver operating characteristic (ROC) curves were

assessed using the "survival" package and "survivalROC" package of R language to evaluate the predictive power of the signal for survival prediction. The "heatmap" package was used for acquiring related graphics such as heatmaps, the distribution of risk scores with survival, and statuses of patients dependent on selected immune LncRNA expression. The patients in the training cohort were divided into a high-risk group and a low-risk group based on the mean risk score. The survival differences between the high-risk group and the low-risk group were evaluated by "Kaplan–Meier" and the "survival" package log-rank test (Wei et al., 2019). The method was also applied to the verification in the validation cohort and the whole cohort.

Principal component analysis (PCA) was used to visualize and compare the predictive power of selected immune LncRNAs and all immune LncRNAs and immune genes. Gene Set Enrichment Analysis (GSEA) was employed to predefine the enrichment of immune gene sets and find different functional phenotypes. All analyses were performed by using the R programming language.[3]

Statistical Analysis

The chi-square test and Student's t-test were adopted to confirm that there is no significant difference in clinical characteristics between the training cohort and validation cohort. Kaplan–Meier survival analysis was performed to compare the survival rate of the high-risk group and low-risk group. The independent prognostic factors in HCC were determined using univariate Cox regression and multivariate Cox regression. SPSS26.0 was used for all statistical analyses. A P value <0.05 is considered significant.

RESULTS

Expression of LncRNAs in HCC Patients

Based on the screening criteria, 319 patients in TCGA-LIHC were selected with matching complete gene expression. We randomly divided them into a training cohort ($n = 160$) and a validation cohort ($n = 159$). The chi-square test showed no significant difference between the two groups (**Table 1**).

We screened out the non-coding RNAs by determining whether the genes in the mRNA matrix encode proteins or not and further sorted out the LncRNA matrix. The correlation test was used between LncRNAs and immune genes, the screening coefficient was set to Cor-Filter >0.4 or <−0.4, while the P value was set to less than 0.001 (Zhang Y. et al., 2019). The non-fluctuation LncRNA expression was deleted in the process. We ended up with 3094 immune-related LncRNAs in the training cohort. Most of the genes were significantly upregulated ($n = 3038$) and a small part of the immune-related LncRNAs were downregulated.

Construction of a Prognostic Model for Survival Prediction in the Training Cohort

Twenty-four LncRNAs were found to be significantly associated with OS in the training cohort by univariate Cox regression.

Least Absolute Shrinkage and Selection Operator Cox regression analysis was performed to identify immune-related LncRNA related to prognosis. The eight most relevant prognostic parameters were calculated and cross-validated to facilitate parameter selection. Some parameters were removed if their coefficient was 0 to prevent overfitting (Liu et al., 2020). After excluding LncRNAs with a coefficient less than 0.1, four immune LncRNAs with the closest relationship to prognosis were selected and a risk score formula was developed through linear combination of the expression values of the four LncRNAs adjusted by the LASSO regression coefficient. The risk score is equal to (expression value of AL603839.3*0.15195028612722) + (expression value of MSC-AS1*0.101019611929165) + (expression value of AL031985.3*0.758428456611441) + (expression value of THUMPD3-AS1*0.127206488953884). The risk score for each patient was calculated based on this equation (**Table 2**).

The median risk score in the training cohort was recorded as the dividing point that is used to divide the cohort into a high-risk group and a low-risk group. Kaplan–Meier analysis showed that the OS in high-risk patients was significantly lower than that in low-risk patients (P value = 3.272e-04). The median survival time was 564 days and 1145 days in each group, respectively. The 5-year survival rates in the high-risk and low-risk groups were 41.5 and 70.5%, respectively (**Figure 1C**). The specific risk score

TABLE 1 | Clinical factors of patients with ICC.

	Training cohort (*n* = 160)	Validation cohort (*n* = 159)	Whole cohort (*n* = 319)	*P* value
Clinical factors				
Gender				
Male	49	51	100	
Female	111	108	219	0.810
Age (years)				
≤60	80	80	160	
>60	80	79	159	1.000
Grade				
G1 + G2	98	100	198	
G3 + G4	62	59	121	0.818
TNM stage				
I + II	120	116	236	
III + IV	40	43	83	0.703
Survival status				
Alive	114	101	215	
Dead	46	58	104	0.153

TABLE 2 | The four immune LncRNAs screened out by univariate Cox regression and LASSO regression analysis.

LncRNAs	Hazard ratio (95% CI)	*P* value	Coefficient
AL603839.3	2.253 (1.436-3.536)	<0.001	0.151950286127222
MSC-AS1	2.088 (1.414-3.083)	<0.001	0.101019611929165
AL031985.3	7.225 (3.068-17.014)	<0.001	0.758428456611441
THUMPD3-AS1	2.861 (1.662-4.924)	<0.001	0.758428456611441

distribution, survival status of the patients, and the selected four prognostic expression profiles also showed that a high expression of the four immune-related LncRNAs correlated with a poor prognosis (**Figure 1B**).

Comparison Between Four LncRNAs Signature and Clinical Characteristics

Multivariate Cox regression analysis showed that the risk score was associated with OS in the training cohort after adjusting for age, gender, grade stage, and Tumor, Node, metastasis (TNM) stage [HR = 4.863, 95% CI (3.004–7.872), P value < 0.001] (Ochenduszko et al., 2017). Although we found that the TNM stage also can evaluate the prognosis as an independent factor significantly in multivariate regression, the AUC of the ROC analysis predicting 5-year OS showed that the risk score was highest at 0.736, indicating that it is more accurate than other traditional clinical parameters in survival prediction (**Figure 1A**).

Testing in the Validation Cohort and the Whole Cohort

The risk score was confirmed to be an independent prognostic factor in the training cohort. Further verification was done in the validation cohort and the whole cohort to confirm the usability of the model. The expression value of immune-related LncRNAs of patients in the validation cohort was used to calculate the risk score depending on the equation. The median value of the risk score in the training cohort (risk score = 0.728) was adopted to divide the validation cohort into a low-risk group (*n* = 78) and a high-risk group (*n* = 81). First of all, univariate analysis and multivariate analysis showed that the risk score correlated with OS after adjustment for age, gender, grade, and TNM stage (**Table 3**; Scheiner et al., 2019). Consistently with this finding, Kaplan-Meier analysis revealed that the OS in high-risk patients was significantly lower than in low-risk patients (P value = 3.66e-03, **Figure 2B**). The 5-year survival rates in the high-risk and low-risk groups were 33.0 and 49.5%, respectively. The specific risk score distribution, survival status of the patients, and the selected four prognostic expression profiles also showed the same result as in the training cohort (**Figure 2C**). The ROC analysis indicated that the risk score is the best signature (AUC = 0.727) among the clinical characteristics (**Figure 2A**).
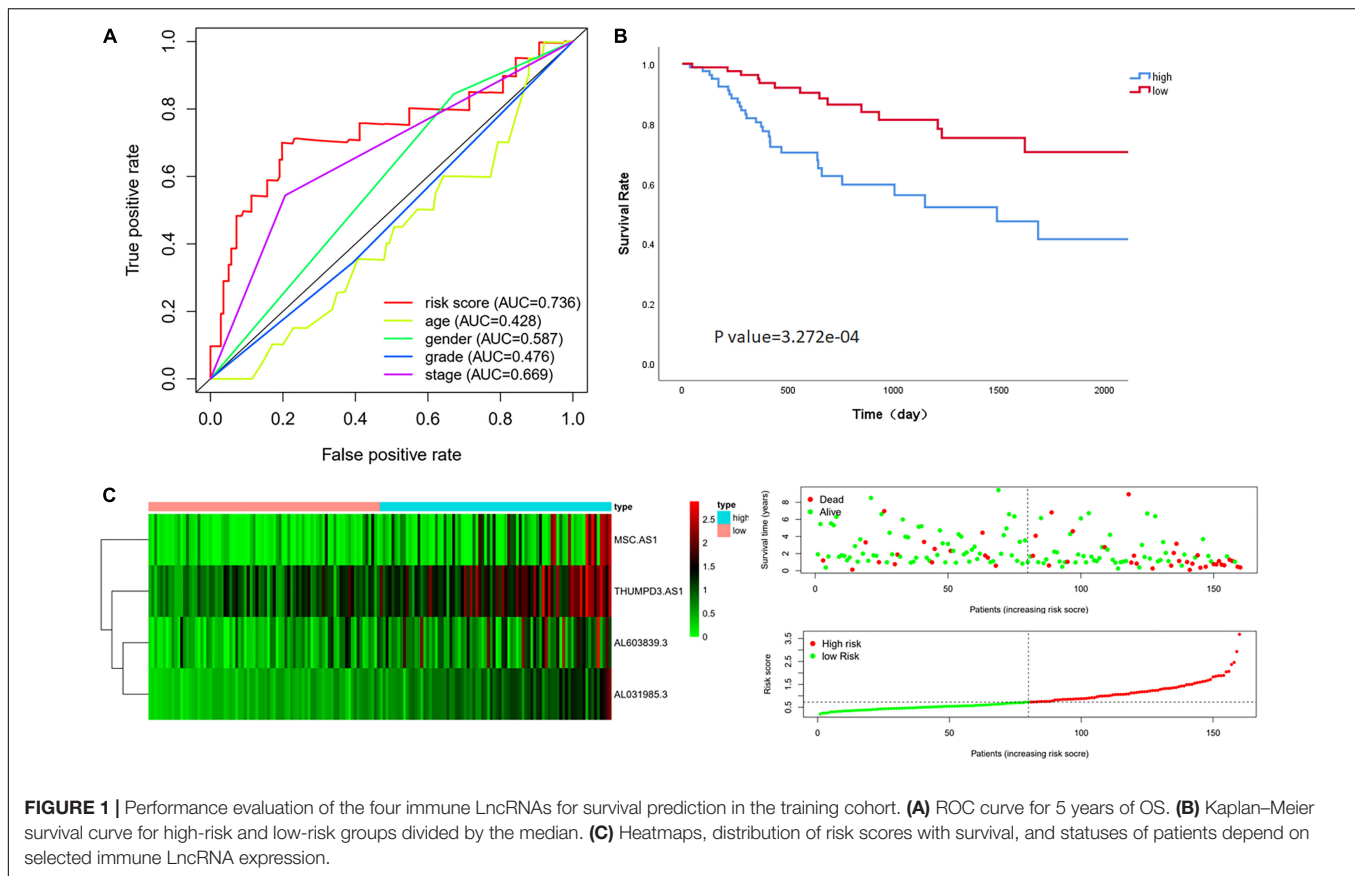
We then further verified the performance of the four immune-related LncRNAs signature in the whole cohort and achieved similar results. The median value of the risk score in the training cohort was used to divide the whole cohort into a low-risk group (*n* = 158) and a high-risk group (*n* = 161). Consistently with this finding, Kaplan–Meier analysis showed that the OS in high-risk patients was significantly lower than that in low-risk patients (P value = 7.184e-06, **Figure 3B**). The five-year survival rates in the high-risk and low-risk groups were 41.8 and 62.3%, respectively. The specific risk score distribution, survival status of the patients, and the selected four prognostic expression profiles of LncRNAs were similar to the result for the other two cohorts (**Figures 3A,C**).

FIGURE 1 | Performance evaluation of the four immune LncRNAs for survival prediction in the training cohort. **(A)** ROC curve for 5 years of OS. **(B)** Kaplan–Meier survival curve for high-risk and low-risk groups divided by the median. **(C)** Heatmaps, distribution of risk scores with survival, and statuses of patients depend on selected immune LncRNA expression.

We compared the effectiveness of our four immune-related LncRNAs signature against two other published LncRNA signatures in HCC by using the ROC curve. The ROC analysis indicated that the risk scores calculated by their signatures were lower than ours and that the AUC values were 0.585 and 0.707 (**Supplementary Figure 1**). This comparison showed that our signature offers an advantage in assessing a prognosis among the clinical characteristics (Yan et al., 2019; Zhang et al., 2020).
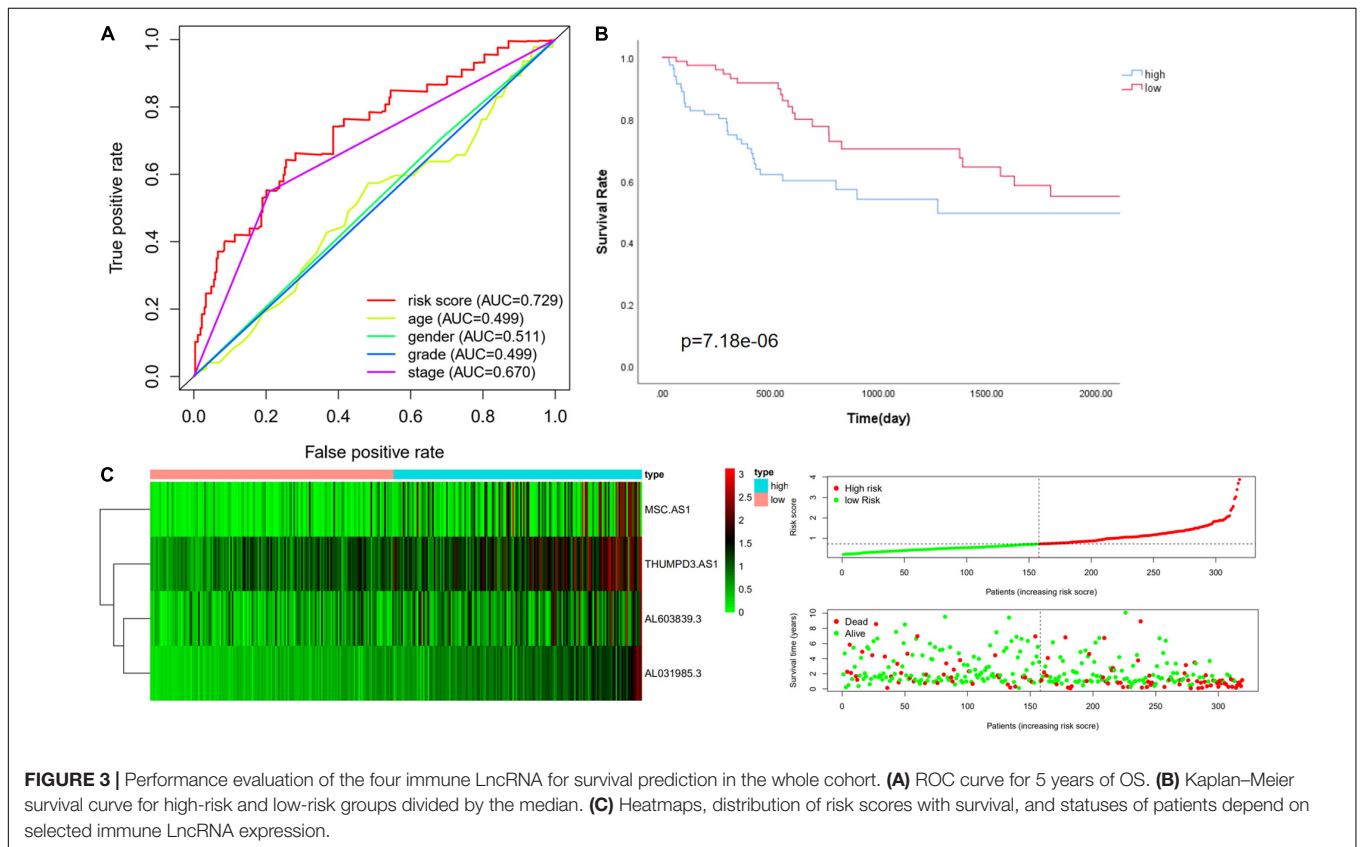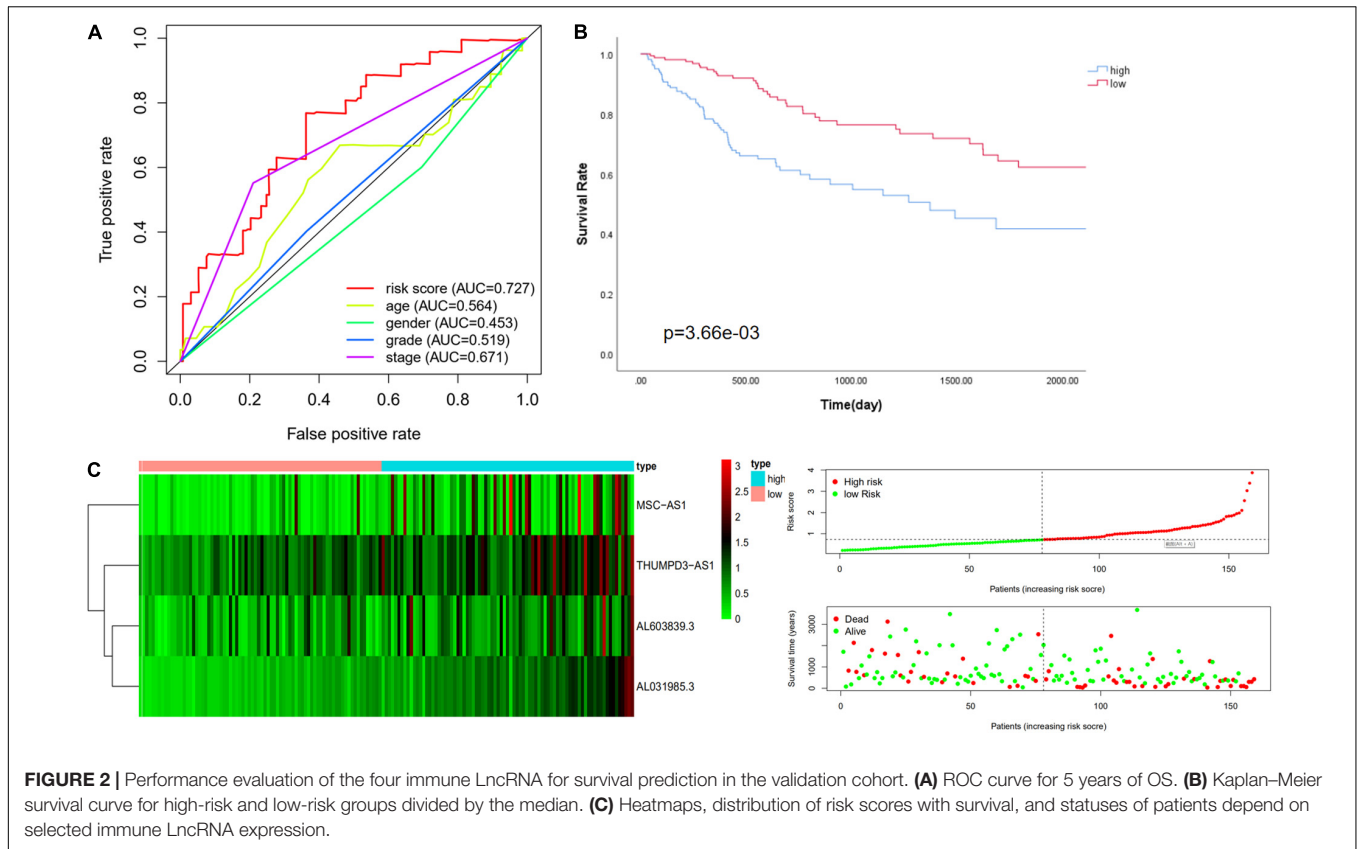
GSEA AND PCA IN THE WHOLE COHORT

In order to verify the significance of immune-related LncRNAs as a label to predict the prognosis of HCC patients, we performed a PCA among selected immune-related LncRNAs (Wei et al., 2019), all genes, immune-related genes, and all immune-related LncRNAs. The PCA analysis suggested that the samples screened by the four immune-related LncRNAs could clearly divide the whole group into a low-risk and high-risk group. In addition, its effect was much better than that of other signatures based on all genes, immune-related genes, or immune-related LncRNAs (**Figure 4**).
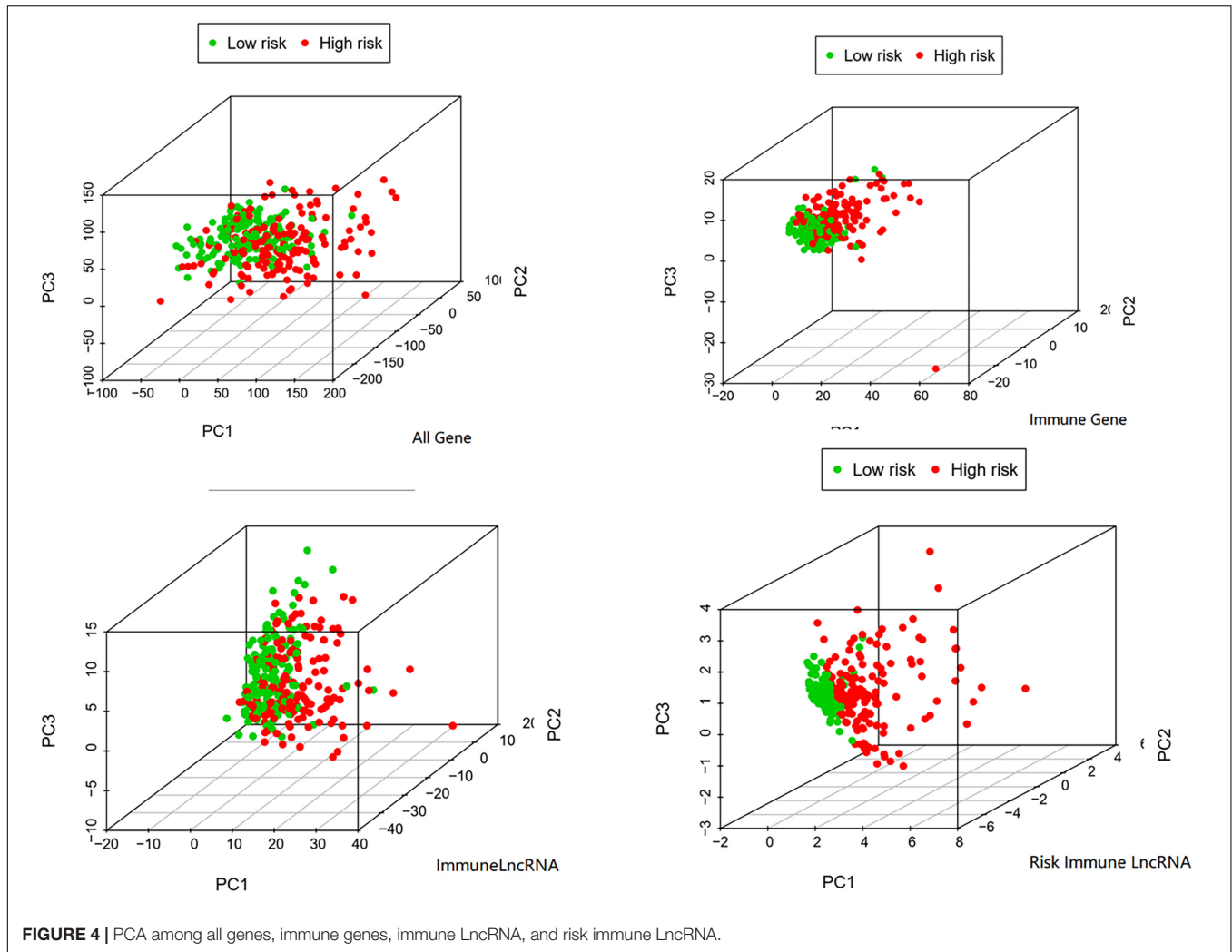
Gene Set Enrichment Analysis showed a higher expression of immune LncRNA obtained from the IMMUNE_RESPONSE gene set and IMMUNE_SYSTEM_PROCESS gene set in high-risk patients, which confirmed that a high expression of relative

immune-related LncRNAs may be positively associated with a poor prognosis, while the rank in the ordered gene set also implied this result (**Figure 5A**). GO and KEGG analysis indicated that the four immune-related LncRNAs might be connected with several biological characteristics including acetylation, inflammatory response, protein binding, stimulatory C-type

TABLE 3 | Univariate and multivariate Cox regression analysis in each cohort.

Variables	Univariate analysis		Multivariate analysis	
	HR (95% CI)	P value	HR (95% CI)	P value
Training cohort (n = 160)				
Age (>60/≤60)	1.007 (0.984–1030)	0.560	1.000 (0.976–1.025)	0.998
Gender (female/male)	1.202 (0.629–2.296)	0.577	1.491 (0.745–2.982)	0.259
Grade (G1 + G2/G3 + G4)	1.015 (0.558–1.847)	0.961	0.717 (0.374–1.374)	0.316
TNM stage (S1 + S2/S3 + S4)	3.332 (1.851–5.999)	<0.001	3.733 (2.005–6.953)	<0.001
Risk score (low/high)	4.519 (2.844–7.182)	<0.001	4.863 (3.004–7.872)	<0.001
Validation cohort (n = 159)				
Age (>60/≤60)	1.007 (0.984–1.030)	0.560	0.997 (0.974–1.021)	0.805
Gender (female/male)	1.202 (0.629–2.296)	0.577	1.008 (0.542–20185)	0.813
Grade (G1 + G2/G3 + G4)	1.015 (0.558–1.847)	0.961	0.699 (0.361–1.352)	0.288
TNM stage (S1 + S2/S3 + S4)	3.332 (1.851–5.999)	<0.001	3.365 (1.808–6.263)	<0.001
Risk score (low/high)	3.627 (2.510–5.241)	<0.001	3.818 (2.565–5.679)	<0.001
Whole cohort (n = 319)				
Age (>60/≤60)	1.005 (0.990–1.020)	0.542	1.004 (0.988–1.020)	0.630
Gender (female/male)	0.815 (0.545–1.218)	0.318	0.788 (0.518–1.200)	0.268
Grade (G1 + G2/G3 + G4)	1.083 (0.727–1.614)	0.694	0.867 (0.568–1.323)	0.508
TNM stage (S1 + S2/S3 + S4)	2.842 (1.925–4.196)	<0.001	2.770 (1.864–4.115)	<0.001
Risk score (low/high)	2.680 (2.089–3.439)	<0.001	2.736 (2.107–3.551)	<0.001

FIGURE 2 | Performance evaluation of the four immune LncRNA for survival prediction in the validation cohort. **(A)** ROC curve for 5 years of OS. **(B)** Kaplan–Meier survival curve for high-risk and low-risk groups divided by the median. **(C)** Heatmaps, distribution of risk scores with survival, and statuses of patients depend on selected immune LncRNA expression.

FIGURE 3 | Performance evaluation of the four immune LncRNA for survival prediction in the whole cohort. **(A)** ROC curve for 5 years of OS. **(B)** Kaplan–Meier survival curve for high-risk and low-risk groups divided by the median. **(C)** Heatmaps, distribution of risk scores with survival, and statuses of patients depend on selected immune LncRNA expression.

FIGURE 4 | PCA among all genes, immune genes, immune LncRNA, and risk immune LncRNA.
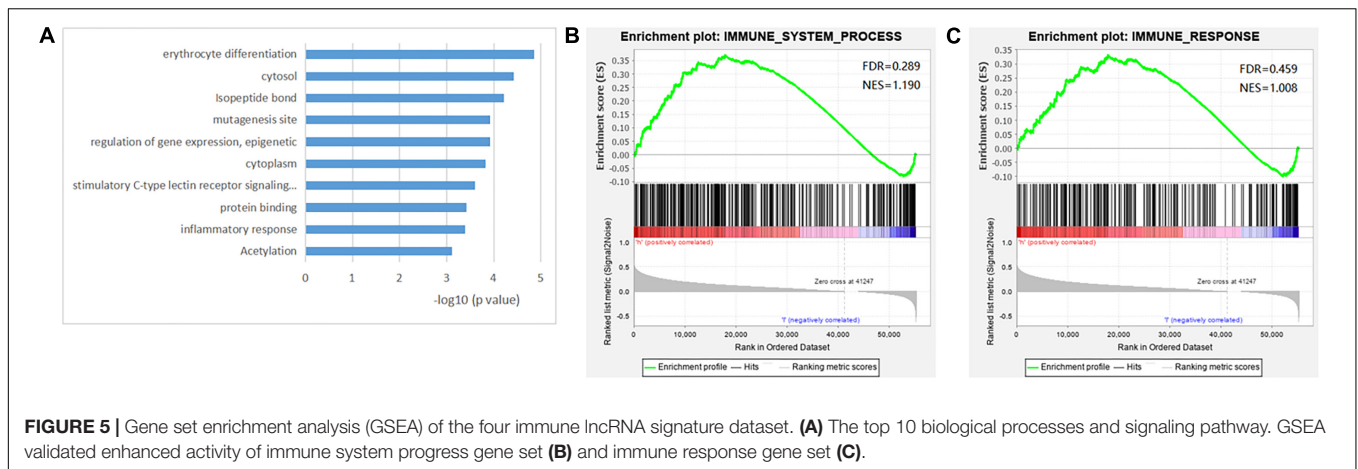
lectin receptor signaling pathway, cytoplasm, regulation of gene expression, epigenetics, mutagenesis site, isopeptide bond, cytosol, and erythrocyte differentiation through the DAVID web annotation tool (**Figures 5B,C**; Shao et al., 2018; Yuan et al., 2018). The analysis could help us to understand the potential molecular mechanisms of the four LncRNAs in the occurrence and development of HCC (Ding et al., 2020).

DISCUSSION

Hepatocellular carcinoma is a well-known malignant tumor with a high morbidity and mortality. Therapy of HCC is still unsatisfactory despite the fact that great progress has been made in the technology of surgery and adjuvant therapies. Finding a significant marker that can more accurately predict the HCC prognosis will provide reliable guidance for clinical treatment. Many studies have documented that the expression of some LncRNAs is related to the prognosis of patients with various malignant tumors, including HCC. In addition, they could be used for the prediction of the prognosis of

patients with better results than traditional clinical parameters (Li et al., 2017; Xie et al., 2018; Zhu et al., 2019). A single LncRNA-based prediction may lead to bias. Therefore, some researchers use comprehensive LncRNA screening to establish an evaluation equation to achieve the desired outcome (Gu et al., 2019). At the same time, studies show that immune reactions play an important role in the occurrence and development of HCC (Hou et al., 2018). We found that the immune-related LncRNAs as a signature were better at predicting the prognosis than a single LncRNA. Additionally, both PCA and GSEA proved that the immune-related LncRNA signature has a better predictive value. We compared the effectiveness of our signature with a published signature by using the ROC curve, which also supports this view (Yan et al., 2019). Our signature showed superiority in comparison with another published immune-related LncRNAs signature, considering that different screening criteria of LncRNA produce different outcomes.

The whole sample was randomly divided into a training cohort and a validation cohort. We then created the whole cohort for further validation. First, we identified 3094 immune-related

FIGURE 5 | Gene set enrichment analysis (GSEA) of the four immune lncRNA signature dataset. **(A)** The top 10 biological processes and signaling pathway. GSEA validated enhanced activity of immune system progress gene set **(B)** and immune response gene set **(C)**.

LncRNAs by matching the gene ID in the IMMUNE_RESPONSE gene set and the IMMUNE_SYSTEM_PROCESS gene set obtained from the GSEA website. Single-factor Cox regression analysis screened 24 survival-related immune LncRNAs and LASSO regression analysis further identified the four most suitable immune-related LncRNAs (AL603839.3, MSC-AS1, AL031985.3, and THUMPD3-AS1) and coefficients. The risk score was used to divide patients into a low-risk group and a high-risk group. It was calculated using the coefficients and the expression of the selected immune-related LncRNAs. The AUC in the ROC of the 5-year survival rate predicted by the risk score in the validation cohort and the whole cohort was 0.727 and 0.729, respectively, which indicated that the model performed better in survival prediction than other conventional clinical parameters. Although the TNM stage also could be used as an independent factor to predict the prognosis through multivariate regression (Andreou et al., 2013), the risk score had a higher AUC and Hazard Ratio.

We then performed the GSEA to explore the role of immune-related LncRNAs in tumor behavior. The analysis showed that the four immune-related LncRNAs were significantly enriched in some biological processes and signaling pathways including acetylation, inflammatory response, protein binding, stimulatory C-type lectin receptor signaling pathway, cytoplasm, regulation of gene expression, epigenetics, mutagenesis site, isopeptide bond, cytosol, and erythrocyte differentiation (Zhou et al., 2017, 2020). It was implied that the four selected LncRNAs might be connected to some gene recombination and cell injury, thus affecting the progression and malignancy of the tumor. Studies showed that gene recombination plays an important role in eliminating the autoreactive B-cell antigen receptor, which is closely related to liver regeneration and liver failure (Behnke et al., 2018; Setz et al., 2019). Damage to the liver caused by surgery or inflammation can lead to regeneration of liver cells and may increase the risk of gene mutations, which are also closely related to the development of tumors (Yin et al., 2011; Suzuki et al., 2016). The findings are likely to be applied to the development of new targeted anti-cancer therapies if the hypothesis can be proved.

This study has several limitations. First, the equation would be more convincing if it could be verified in other databases. Actually, the LncRNAs obtained from the TCGA may not match up with those in other databases because of the use of different chips and recording methods. In other datasets that we have consulted, such as SEER, ICGC, and TANRIC, we could not find a dataset that has both the clinical data and corresponding expression of the LncRNAs. It would make the results more convincing if the effectiveness of our signature could be validated in other datasets. Second, we only performed enrichment analysis and made assumptions regarding the function of these four immune-related LncRNAs without additional research to further explore the mechanism. Third, in vitro experiments also could make the results more convincing. However, the sufficient survival time cannot be collected in a short period of time because the testing for expression of LncRNA requires fresh tissue samples. The further analysis of bio-information results combined with clinical data is also the direction that our follow-up research will take. Finally, although the total sample size was 319, a larger sample would produce more convincing results.

In summary, we developed a four-immune-related-LncRNA signature for the prediction of HCC prognosis after surgery. In addition, this signature might guide the use of immunotherapy in HCC.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study, these can be found in The Cancer Genome Atlas (https://portal.gdc.cancer.gov/).

ETHICS STATEMENT

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. Written informed consent for participation was not required for this study in accordance with the national legislation and the institutional requirements.

AUTHOR CONTRIBUTIONS

MuL and MiL: manuscript writing including the figures and tables. TW and ZC: guidance and design in figures and tables. XW and WX: manuscript writing. SS and BP: conception and design and manuscript writing. All authors: final approval of the manuscript.

REFERENCES

Ahmad, I., Valverde, A., Ahmad, F., and Naqvi, A. R. (2020). Long noncoding RNA in myeloid and lymphoid cell differentiation, polarization and function. *Cells* 9:269. doi: 10.3390/cells9020269

Andreou, A., Vauthey, J. N., Cherqui, D., Zimmitti, G., Ribero, D., Truty, M. J., et al. (2013). Improved long-term survival after major resection for hepatocellular carcinoma: a multicenter analysis based on a new definition of major hepatectomy. *J. Gastrointest. Surg.* 17, 66–77. discussion.77 doi: 10.1007/s11605-012-2005-4

Behnke, K., Zhuang, Y., Xu, H. C., Sundaram, B., Reich, M., Shinde, P. V., et al. (2018). B cell-mediated maintenance of cluster of differentiation 169-positive cells is critical for liver regeneration. *Hepatology* 68, 2348–2361. doi: 10.1002/hep.30088

Bhan, A., Soleimani, M., and Mandal, S. S. (2017). Long noncoding RNA and cancer: a new paradigm. *Cancer Res.* 77, 3965–3981. doi: 10.1158/0008-5472.can-16-2634

Ding, X. B., Jin, J., Tao, Y. T., Guo, W. P., Ruan, L., Yang, Q. L., et al. (2020). Predicted *Drosophila* interactome resource and web tool for functional interpretation of differentially expressed genes. *Database J. Biol. Databases Curat.* 2020:baaa005.

European Association for the Study of the Liver, Electronic address, easloffice easloffice eu., and European Association for the Study of the Liver (2018). EASL clinical practice guidelines: management of hepatocellular carcinoma. *J. Hepatol.* 69, 182–236.

Fife, B. T., Pauken, K. E., Eagar, T. N., Obu, T., Wu, J., Tang, Q., et al. (2009). Interactions between PD-1 and PD-L1 promote tolerance by blocking the TCR-induced stop signal. *Nat. Immunol.* 10, 1185–1192. doi: 10.1038/ni.1790

Francisco, L. M., Salinas, V. H., Brown, K. E., Vanguri, V. K., Freeman, G. J., Kuchroo, V. K., et al. (2009). PD-L1 regulates the development, maintenance, and function of induced regulatory T cells. *J. Exp. Med.* 206, 3015–3029. doi: 10.1084/jem.20090847

Gao, J., Aksoy, B. A., Dogrusoz, U., Dresdner, G., Gross, B., Sumer, S. O., et al. (2013). Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci. Signal.* 6:pl1. doi: 10.1126/scisignal.2004088

Gu, J. X., Zhang, X., Miao, R. C., Xiang, X. H., Fu, Y. N., Zhang, J. Y., et al. (2019). Six-long non-coding RNA signature predicts recurrence-free survival in hepatocellular carcinoma. *World J. Gastroenterol.* 25, 220–232. doi: 10.3748/wjg.v25.i2.220

Gupta, S., Lee, R. E. C., and Faeder, J. R. (2020). Parallel tempering with lasso for model reduction in systems biology. *PLoS Comput. Biol.* 16:e1007669. doi: 10.1371/journal.pcbi.1007669

He, Y., Meng, X. M., Huang, C., Wu, B. M., Zhang, L., Lv, X. W., et al. (2014). Long noncoding RNAs: novel insights into hepatocelluar carcinoma. *Cancer Lett.* 344, 20–27. doi: 10.1016/j.canlet.2013.10.021

Hilmi, M., Vienot, A., Rousseau, B., and Neuzillet, C. (2019). Immune therapy for liver cancers. *Cancers* 12, 7–34.

Hou, X. J., Ye, F., Li, X. Y., Liu, W. T., Jing, Y. Y., Han, Z. P., et al. (2018). Immune response involved in liver damage and the activation of hepatic progenitor cells during liver tumorigenesis. *Cell. Immunol.* 326, 52–59. doi: 10.1016/j.cellimm.2017.08.004

Jia, Y., Zeng, Z., Li, Y., Li, Z., Jin, L., Zhang, Z., et al. (2015). Impaired function of CD4+ T follicular helper (Tfh) cells associated with hepatocellular carcinoma progression. *PLoS One* 10:e0117458. doi: 10.1371/journal.pone.0117458

Jing, H., Qu, X., Liu, L., and Xia, H. (2018). A novel long noncoding RNA (lncRNA), LL22NC03-N64E9.1, promotes the proliferation of lung cancer cells and is a potential prognostic molecular biomarker for lung cancer. *Med. Sci. Monit.* 24, 4317–4323. doi: 10.12659/msm.908359

Klec, C., Prinz, F., and Pichler, M. (2019). Involvement of the long noncoding RNA NEAT1 in carcinogenesis. *Mol. Oncol.* 13, 46–60. doi: 10.1002/1878-0261.12404

Kubes, P., and Jenne, C. (2018). Immune responses in the liver. *Annu. Rev. Immunol.* 36, 247–277.

Li, J., Han, L., Roebuck, P., Diao, L., Liu, L., Yuan, Y., et al. (2015). TANRIC: an interactive open platform to explore the function of lncRNAs in Cancer. *Cancer Res.* 75, 3728–3737. doi: 10.1158/0008-5472.can-15-0273

Li, J. K., Chen, C., Liu, J. Y., Shi, J. Z., Liu, S. P., Liu, B., et al. (2017). Long noncoding RNA MRCCAT1 promotes metastasis of clear cell renal cell carcinoma via inhibiting NPR3 and activating p38-MAPK signaling. *Mol. Cancer* 16:111.

Li, Y., Jiang, T., Zhou, W., Li, J., Li, X., Wang, Q., et al. (2020). Pan-cancer characterization of immune-related lncRNAs identifies potential oncogenic biomarkers. *Nat. Commu.* 11:1000.

Liu, Y., Wang, L., Liu, H., Li, C., and He, J. (2020). The prognostic significance of metabolic syndrome and a related Six-lncRNA signature in esophageal squamous cell carcinoma. *Front. Oncol.* 10:61. doi: 10.3389/fonc.2020.00061

Llovet, J. M., Kudo, M., Cheng, A.-L., Finn, R. S., Galle, P. R., Kaneko, S., et al. (2019). Lenvatinib (len) plus pembrolizumab (pembro) for the first-line treatment of patients (pts) with advanced hepatocellular carcinoma (HCC): phase 3 LEAP-002 study. *J. Clin. Oncol.* 3:TS4152.

Ma, Z., Huang, H., Xu, Y., He, X., Wang, J., Hui, B., et al. (2017). Current advances of long non-coding RNA highly upregulated in liver cancer in human tumors. *OncoTargets Ther.* 10, 4711–4717. doi: 10.2147/ott.s136915

Mizukoshi, E., Nakamoto, Y., Arai, K., Yamashita, T., Sakai, A., Sakai, Y., et al. (2011). Comparative analysis of various tumor-associated antigen-specific t-cell responses in patients with hepatocellular carcinoma. *Hepatology* 53, 1206–1216. doi: 10.1002/hep.24149

Nanni, S., Bacci, L., Aiello, A., Re, A., Salis, C., Grassi, C., et al. (2020). Signaling through estrogen receptors modulates long non-coding RNAs in prostate cancer. *Mol. Cell. Endocrinol.* 511:110864. doi: 10.1016/j.mce.2020.110864

Ochenduszko, S., Puskulluoglu, M., Konopka, K., Fijorek, K., Slowik, A. J., Pêdziwiatr, M., et al. (2017). Clinical effectiveness and toxicity of second-line irinotecan in advanced gastric and gastroesophageal junction adenocarcinoma: a single-center observational study. *Ther. Adv. Med. Oncol.* 9, 223–233. doi: 10.1177/1758834016689029

Patel, S. P., and Kurzrock, R. (2015). PD-L1 expression as a predictive biomarker in cancer immunotherapy. *Mol. Cancer Ther.* 14, 847–856. doi: 10.1158/1535-7163.mct-14-0983

Sangro, B., Gomez-Martin, C., de la Mata, M., Iñarrairaegui, M., Garralda, E., Barrera, P., et al. (2013). A clinical trial of CTLA-4 blockade with tremelimumab

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2020.566491/full#supplementary-material

Supplementary Figure 1 | The effectiveness of our signature with a published signature by using the ROC curve, A (PMID:30206981) and B (PMID:32330311).

in patients with hepatocellular carcinoma and chronic hepatitis C. *J. Hepatol.* 59, 81–88. doi: 10.1016/j.jhep.2013.02.022

Scheiner, B., Kirstein, M., Popp, S., Hucke, F., Bota, S., Rohr-Udilova, N., et al. (2019). Association of platelet count and mean platelet volume with overall survival in patients with cirrhosis and unresectable hepatocellular carcinoma. *Liver Cancer* 8, 203–217. doi: 10.1159/000489833

Setz, C. S., Khadour, A., Renna, V., Iype, J., Gentner, E., He, X., et al. (2019). Pten controls B-cell responsiveness and germinal center reaction by regulating the expression of IgD BCR. *EMBO J.* 38:e100249.

Shao, X., Bao, W., Hong, X., Jiang, H., and Yu, Z. (2018). Identification and functional analysis of differentially expressed genes associated with cerebral ischemia/reperfusion injury through bioinformatics methods. *Mol. Med. Rep.* 18, 1513–1523.

Shen, Y., Peng, X., and Shen, C. (2020). Identification and validation of immune-related lncRNA prognostic signature for breast cancer. *Genomics* 112, 2640–2646. doi: 10.1016/j.ygeno.2020.02.015

Shivaji, U. N., Jeffery, L., Gui, X., Smith, S. C. L., Ahmad, O. F., Akbar, A., et al. (2019). Immune checkpoint inhibitor-associated gastrointestinal and hepatic adverse events and their management. *Ther. Adv. Gastroenterol.* 12:1756284819884196.

Siegel, R. L., Miller, K. D., and Jemal, A. (2019). Cancer statistics, 2019. *CA Cancer J. Clin.* 69, 7–34.

Sprinzl, M. F., and Galle, P. R. (2017). Current progress in immunotherapy of hepatocellular carcinoma. *J. Hepatol.* 66, 482–484. doi: 10.1016/j.jhep.2016.12.009

Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15545–15550. doi: 10.1073/pnas.0506580102

Sun, J., Zhang, Z., Bao, S., Yan, C., Hou, P., Wu, N., et al. (2020). Identification of tumor immune infiltration-associated lncRNAs for improving prognosis and immunotherapy response of patients with non-small cell lung cancer. *J. Immunother. Cancer* 8:e000110. doi: 10.1136/jitc-2019-000110

Suzuki, Y., Katagiri, H., Wang, T., Kakisaka, K., Kume, K., Nishizuka, S. S., et al. (2016). Ductular reactions in the liver regeneration process with local inflammation after physical partial hepatectomy. *Lab. Invest.* 96, 1211–1222. doi: 10.1038/labinvest.2016.97

Teng, M. W., Galon, J., Fridman, W. H., and Smyth, M. J. (2015). From mice to humans: developments in cancerimmunoediting. *J. Clin. Invest.* 125, 3338–3346. doi: 10.1172/jci80004

Wang, W., Zhao, Z., Yang, F., Wang, H., Wu, F., Liang, T., et al. (2018). An immune-related lncRNA signature for patients with anaplastic gliomas. *J. Neuro Oncol.* 136, 263–271. doi: 10.1007/s11060-017-2667-6

Wang, Y., Zhong, H., Xie, X., Chen, C. Y., Huang, D., Shen, L., et al. (2015). Long noncoding RNA derived from CD244 signaling epigenetically controls CD8+ T-cell immune responses in tuberculosis infection. *Proc. Natl. Acad. Sci. U. S. A.* 112, E3883–E3892.

Wei, C., Liang, Q., Li, X., Li, H., Liu, Y., Huang, X., et al. (2019). Bioinformatics profiling utilized a nine immune-related long noncoding RNA signature as a prognostic target for pancreatic cancer. *J. Cell Biochem.* 120, 14916–14927. doi: 10.1002/jcb.28754

Xie, Y., Zhang, Y., Du, L., Jiang, X., Yan, S., Duan, W., et al. (2018). Circulating long noncoding RNA act as potential novel biomarkers for diagnosis and prognosis of non-small cell lung cancer. *Mol. Oncol.* 12, 648–658. doi: 10.1002/1878-0261.12188

Yan, J., Zhou, C., Guo, K., Li, Q., and Wang, Z. (2019). A novel seven-lncRNA signature for prognosis prediction in hepatocellular carcinoma. *J. Cell Biochem.* 120, 213–223. doi: 10.1002/jcb.27321

Yin, S., Wang, H., Park, O., Wei, W., Shen, J., and Gao, B. (2011). Enhanced Liver regeneration in IL-10-deficient mice after partial hepatectomy via stimulating inflammatory response and activating hepatocyte STAT3. *Am. J. Pathol.* 178, 1614–1621. doi: 10.1016/j.ajpath.2011.01.001

Yn, W., Yu, C. Y., and Jin, H. Z. (2020). RNA N-methyladenosine modifications and the immune response. *J. Immunol. Res.* 2020:6327614.

Yu, W. D., Wang, H., He, Q. F., Xu, Y., and Wang, X. C. (2018). Long noncoding RNAs in cancer-immunity cycle. *J. Cell Physiol.* 233, 6518–6523. doi: 10.1002/jcp.26568

Yuan, F., Lu, L., Zhang, Y., Wang, S., and Cai, Y. D. (2018). Data mining of the cancer-related lncRNAs GO terms and KEGG pathways by using mRMR method. *Math. Biosci.* 304, 1–8. doi: 10.1016/j.mbs.2018.08.001

Zhang, M., Zhu, K., Pu, H., Wang, Z., Zhao, H., Zhang, J., et al. (2019). An immune-related signature predicts survival in patients with lung adenocarcinoma. *Front. Oncol.* 9:1314. doi: 10.3389/fonc.2019.01314

Zhang, X., Sun, S., Pu, J. K., Tsang, A. C., Lee, D., Man, V. O., et al. (2012). Long non-coding RNA expression profiles predict clinical phenotypes in glioma. *Neurobiol. Dis.* 48, 1–8. doi: 10.1016/j.nbd.2012.06.004

Zhang, X., Zhou, Y., Chen, C., Fang, W., Cai, X., Zhang, X., et al. (2019). Hepatitis B virus reactivation in cancer patients with positive Hepatitis B surface antigen undergoing PD-1 inhibition. *J. Immunother. Cancer* 7:322.

Zhang, Y., Li, Z., Chen, M., Chen, H., Zhong, Q., Liang, L., et al. (2019). Identification of a new eight-long noncoding RNA molecular signature for breast cancer survival prediction. *DNA Cell Biol.* 38, 1529–1539. doi: 10.1089/dna.2019.5059

Zhang, Y., Zhang, L., Xu, Y., Wu, X., Zhou, Y., and Mo, J. (2020). Immune-related long noncoding RNA signature for predicting survival and immune checkpoint blockade in hepatocellular carcinoma. *J. Cell Physiol.* 235, 9304–9316. doi: 10.1002/jcp.29730

Zheng, Y., Tian, X., Wang, T., Xia, X., Cao, F., Tian, J., et al. (2019). Long noncoding RNA Pvt1 regulates the immunosuppression activity of granulocytic myeloid-derived suppressor cells in tumor-bearing mice. *Mol. Cancer* 18:61.

Zhong, J. H., Du, X. K., Xiang, B. D., and Li, L. Q. (2016). Adjuvant sorafenib in hepatocellular carcinoma: a cautionary comment of STORM trial. *World J. Hepatol.* 8, 957–960. doi: 10.4254/wjh.v8.i23.957

Zhou, M., Zhang, Z., Bao, S., Hou, P., Yan, C., Su, J., et al. (2020). Computational recognition of lncRNA signature of tumor-infiltrating B lymphocytes with potential implications in prognosis and immunotherapy of bladder cancer. *Brief. Bioinform.* 1–13. doi: 10.1007/bf01789124

Zhou, M., Zhao, H., Xu, W., Bao, S., Cheng, L., and Sun, J. (2017). Discovery and validation of immune-associated long non-coding RNA biomarkers associated with clinically molecular subtype and prognosis in diffuse large B cell lymphoma. *Mol. Cancer* 16:16.

Zhu, S., Huang, X., Zhang, K., Tan, W., Lin, Z., He, Q., et al. (2019). Low expression of long noncoding RNA CTC-297N7.9 predicts poor prognosis in patients with hepatocellular carcinoma. *Cancer Med.* 8, 7679–7692. doi: 10.1002/cam4.2618

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

The reviewer WL declared a shared affiliation, though no other collaboration, with the authors.