Innovative signature establishment using lymphangiogenesis-related lncRNA pairs to predict prognosis of hepatocellular carcinoma

Jincheng Cao a,b,1, Yanni Xu a,1, Xiaodi Liu a,b,1, Yan Cai c, Baoming Luo a,*

a Department of Ultrasound, Sun Yat-sen Memorial Hospital, Sun Yat-sen University, Guangzhou, 510120, China
b Guangdong Provincial Key Laboratory of Malignant Tumor Epigenetics and Gene Regulation, Medical Research Center, Sun Yat-Sen Memorial Hospital, Sun Yat-Sen University, Guangzhou 510120, China
c Department of Ultrasound, Central People’s Hospital of Zhanjiang, 236 Yuanzhu Road, Zhanjiang, Guangdong 524045, China

ARTICLE INFO

Keywords:
Lymphangiogenesis-related long noncoding RNAs
Prognostic signature
Hepatocellular carcinoma
Tumoral infiltration of immune cells
Immunocheckpoint
Chemotherapy

ABSTRACT

Aims: Hepatocellular carcinoma (HCC) remains a major tumoral burden globally, and its heterogeneity encumbers prognostic prediction. The lymphangiogenesis-related long non-coding RNAs (lIrncRNAs) reported to be implicated in immune response regulation show potential importance in predicting the prognostic and therapeutic outcome. Hence, this study aims to establish a lIrncRNA pairs-based signature not requiring specific expression levels of transcripts, which displays promising clinical practicality and satisfactory predictive capability.

Main methods: Transcriptomic and clinical information of the Liver Hepatocellular Carcinoma (LIHC) project retrieved from the TCGA portal were used to find differently expressed lIrncRNA (DElIrncRNA) via analysis performed between lymphangiogenesis-related genes (lr-genes) and lncRNAs(lIrncRNA), and to ultimately construct the signature based on lIrncRNA pairs screened out via Lasso and Cox regression analyses. Akaike information criterion (AIC) values were computed to find the cut-off point optimum for high-risk and low-risk group allocation. The signature then underwent trials in terms of its predictive value for survival, clinicopathological features, immune cells infiltration in tumoral microenvironment, selected checkpoint biomarkers and chemosensitivity.

Key findings: A novel lymphangiogenesis-related lncRNA pair signature was established using nine lIrncRNA pairs identified and significantly related to overall survival, clinicopathological features, immune cells infiltration and susceptibility to chemotherapy. Moreover, the signature efficacy was verified in acknowledged clinicopathological subgroups and partially validated by qRT-PCR assay in various human HCC cell lines.

Significance: The novel lIrncRNA-pairs based signature was shown to effectively and independently estimate HCC prognosis and help screen patients suitable for anti-tumor immunotherapy and chemotherapy.

1. Introduction

Despite the rapid advancement of diagnostic and therapeutic techniques, primary liver cancer, of which hepatocellular carcinoma (HCC) constitutes approximately 75%, remains a major constituent to the universal tumor burden (McGlynn et al., 2021). Prevalence of the most meaningful risk factors for HCC at present, hepatitis B and C infection, should drop in years to come owing to vaccination of the newborns (Akinyemiju et al., 2017). However, metabolic risk factors, inclusive of adiposity (Lauby-Secretan et al., 2016), diabetes (Ohkuma et al., 2018), local lymphatic metastasis (Qin and Tang, 2002) and alcoholic abuse (Petrick et al., 2018), are displaying increasing importance and tend to jointly become the major cause of HCC worldwide. This encumbers HCC prognosis prediction due to high heterogeneity of HCC and different risk factors impacting the disease advancement (Torrecilla et al., 2017), thus necessitating the identification of new biomarkers useful for better prognostic prediction and treatment.

Lymphangiogenesis, the process of lymphatic vessel formation, is deeply involved in homeostasis, metabolism and immunity (Suzuki-I-noue et al., 2020). More specifically, tumor-associated lymphangiogenesis plays an essential role in tumor pathogenesis and metastasis through mechanisms like providing niches for tumor stem cells and inhibiting antitumor immune responses (Hu and Luo, 2018). A previous study has revealed that lymphangiogenesis exerted a consequential

* Corresponding author.
E-mail address: luobm@mail.sysu.edu.cn (B. Luo).
1 These authors have contributed equally to this work and share first authorship.

https://doi.org/10.1016/j.heliyon.2022.e10215
Received 10 March 2022; Received in revised form 17 May 2022; Accepted 2 August 2022
2405-8440/© 2022 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
| Gene Symbol | Description | Category | Gifts | GC id               | Relevance score |
|------------|-------------|----------|-------|---------------------|-----------------|
| VEGFC      | Vascular Endothelial Growth Factor C | Protein Coding | 46    | GC04M176683         | 16.69894028     |
| FLT4       | Fms Related Receptor Tyrosine Kinase 4 | Protein Coding | 50    | GC05M186007         | 16.6914579      |
| VEGFD      | Vascular Endothelial Growth Factor D | Protein Coding | 33    | GC0X0M015345        | 13.5329739      |
| CALCRL     | Calcitonin Receptor Like Receptor | Protein Coding | 44    | GC02M187341         | 8.25269223      |
| PDPN       | Podoplanin | Protein Coding | 39    | G01IP13583          | 8.003684351     |
| VEGFA      | Vascular Endothelial Growth Factor A | Protein Coding | 47    | G06P043770          | 7.84228772      |
| LYVE1      | Lymphatic Vessel Endothelial Hyaluronan Receptor 1 | Protein Coding | 41    | G11M101753          | 7.15224541      |
| KDR        | Kinase Insert Domain Receptor | Protein Coding | 52    | G04M050078          | 7.094721794     |
| PROX1      | Prospero Homeobox 1 | Protein Coding | 42    | G01P129883          | 6.58164493      |
| PTN14      | Protein Tyrosine Phosphatase Non-Receptor Type 14 | Protein Coding | 43    | G01M124348          | 6.41293430      |
| SOX18      | SRY-Box Transcription Factor 18 | Protein Coding | 39    | G02M064047          | 4.160199642     |
| FLT1       | Fms Related Receptor Tyrosine Kinase 1 | Protein Coding | 49    | G13M028300          | 3.471974605     |
| CCB1       | Collagen And Calcium Binding EGF Domains 1 | Protein Coding | 40    | G18M059430          | 3.46608689      |
| ANGPT2     | Angiopoietin 2 | Protein Coding | 44    | G08M006499          | 3.38362429      |
| FOXC2      | Forkhead Box C2 | Protein Coding | 44    | G16P086574          | 3.30765765      |
| NRP2       | Neuropilin 2 | Protein Coding | 43    | G02P026581          | 3.15085763      |
| FGFR2      | Fibroblast Growth Factor 2 | Protein Coding | 46    | G04P122826          | 2.56522204      |
| PFG        | Placental Growth Factor | Protein Coding | 42    | G14M074941          | 2.62571947      |
| TGFBI      | Transforming Growth Factor Beta 1 | Protein Coding | 50    | G19M041301          | 2.57847272      |
| CCR7       | C-C Motif Chemokine Receptor 7 | Protein Coding | 44    | G17M040556          | 2.55449558      |
| HIF1A      | Hypoxia Inducible Factor 1 Subunit Alpha | Protein Coding | 46    | G14P061695          | 2.47803287      |
| HMGB1      | High Mobility Group Box 1 | Protein Coding | 44    | G13M030456          | 2.41782546      |
| CXC8       | C-X-C Motif Chemokine Receptor 4 | Protein Coding | 51    | G02M136114          | 2.30786205      |
| POSTN      | Periostin | Protein Coding | 42    | G13M037562          | 2.29275667      |
| CCL21      | C-C Motif Chemokine Ligand 21 | Protein Coding | 41    | G09M034709          | 2.27764933      |
| VASH1      | Vasothesis 1 | Protein Coding | 36    | G14P076761          | 2.17816591      |
| SHH        | Sonic Hedgehog Signaling Molecule | Protein Coding | 48    | G07M155799          | 2.20368274      |
| STAT3      | Signal Transducer And Activator Of Transcription 3 | Protein Coding | 51    | G17M042313          | 2.19051075      |
| ANGPT1     | Angiopoietin 1 | Protein Coding | 45    | G08M107246          | 2.14707403      |
| NRP1       | Neuropilin 1 | Protein Coding | 45    | G10M033177          | 2.16389465      |
| ERBB2      | Erb-B2 Receptor Tyrosine Kinase 2 | Protein Coding | 52    | G17P039687          | 2.15075182      |
| NR2F2      | Nuclear Receptor Subfamily 2 Group F Member 2 | Protein Coding | 48    | G15P096325          | 2.107103348     |
| CXCL12     | C-X-C Motif Chemokine Ligand 12 | Protein Coding | 44    | G10M044294          | 2.0881694       |
| VEGFB      | Vascular Endothelial Growth Factor B | Protein Coding | 43    | G11P064234          | 2.05249567      |
| HPS6       | Heparanase | Protein Coding | 44    | G04M083292          | 2.02933351      |
| TEK        | TEC Receptor Tyrosine Kinase | Protein Coding | 49    | G09P027109          | 2.020673752     |
| HGF        | Hepatocyte Growth Factor | Protein Coding | 50    | G07M081699          | 2.00943422      |
| MET        | MET Proto-Oncogene, Receptor Tyrosine Kinase | Protein Coding | 52    | G07P1116672         | 1.98615053      |
| CEACAM1     | CEA Cell Adhesion Molecule 1 | Protein Coding | 41    | G19M042507          | 1.98037266      |
| NFkB1       | Nuclear Factor kappa B Subunit 1 | Protein Coding | 51    | G04P102501          | 1.93783965      |
| NOS2       | Nitric Oxide Synthase 2 | Protein Coding | 49    | G17M027756          | 1.93575632      |
| S1PR1       | Sphingosine-1-Phosphate Receptor 1 | Protein Coding | 44    | G01P120126          | 1.91859936      |
| FOXC1      | Forkhead Box C1 | Protein Coding | 42    | G06P061610          | 1.90032332      |
| PDGFB       | Platelet Derived Growth Factor Subunit B | Protein Coding | 48    | G22M051477          | 1.8937194       |
| CLEC14A     | C-Type Lectin Domain Containing 14A | Protein Coding | 33    | G14M038254          | 1.842555404     |
| SMAD4       | SMAD Family Member 4 | Protein Coding | 49    | G18P051028          | 1.834085941     |
| IL17A       | Interleukin 17A | Protein Coding | 41    | G06P052186          | 1.829650402     |
| ITGA4       | Integrin Subunit Alpha 4 | Protein Coding | 48    | G02P181456          | 1.812556744     |
| IL7R        | Interleukin 7 Receptor | Protein Coding | 45    | G05P035852          | 1.789921403     |
| TNF         | Tumor Necrosis Factor | Protein Coding | 49    | G06P061170          | 1.782137632     |
| SIX1        | SIX Homeobox 1 | Protein Coding | 43    | G14M060643          | 1.782137632     |
| MMP9        | Matrix Metallopeptidase 9 | Protein Coding | 52    | G20P046008          | 1.76113379      |
| SMARCA4     | SWI/SNF Related, Matrix Associated, Actin Dependent Regulator Of Chromatin, Subfamily A, Member 4 | Protein Coding | 48    | G19P010932          | 1.741836548     |
| IL6         | Interleukin 6 | Protein Coding | 48    | G07P022725          | 1.740508795     |
| MIR27B      | MicroRNA 27b | RNA Gene | 21    | G09P059097          | 1.720370231     |
| MIR9-1      | MicroRNA 9-1 | RNA Gene | 21    | G01M156420          | 1.70369806      |
| PECAM1      | Platelet And Endothelial Cell Adhesion Molecule 1 | Protein Coding | 40    | G17M064319          | 1.695909262     |

(continued on next page)
influence on the survival of HCC patients (Thelen et al., 2009). Some lymphangiogenic genes have been used as biomarkers to predict the prognosis of patients with colorectal liver metastasis after partial hepatectomy (Vellinga et al., 2017). However, prediction based on messenger RNA could suffer from unsatisfactory accuracy due to its inadequate tissue specificity (Deveson et al., 2017). Therefore, it is of necessity to develop new lymphangiogenesis-related biomarkers for better prognostic prediction of HCC.

Long non-coding RNAs (lncRNAs), transcripts whose length is greater than 200 nucleotides, function by regulating gene expression at the post-transcriptional level instead of coding functional proteins (Statello et al., 2021). Accounting for over two-thirds of human transcriptome, lncRNAs were proved to be more sensitive and applicable than the ones including the regulation of lymphatic vasculature (Iyer et al., 2015; Md Yusof et al., 2020). LncRNAs are also active participants in tumorigenesis, as evidenced by their roles in liver cancer axis (Zhao and Lawless, 2017). Some previous evidence has demonstrated that lncRNA-based signatures chosen lncRNAs to be normalized, which is necessary for reducing batch effects among platforms in order to qualify for clinical application (Lv et al., 2020). For example, Ranran Zhou et al. used a ferroptosis-related lncRNA signature consisting of 22 lncRNA pairs to estimate bladder cancer prognostic and immune features (Zhou et al., 2021). However, the number of studies using lncRNA pairs signature for tumoral prognosis prediction is relatively limited. In the present study, we established an innovative prognostic signature using nine lymphangiogenesis-related lncRNA pairs, which efficiently predicted patient survival, tumor infiltration of immune cells, immunomarker gene expression, and chemosensitivity.

2. Materials and methods

2.1. Retrieval of transcriptomic and clinical data

Transcriptomic profiling and clinical information of patients with hepatocellular carcinoma were retrieved from the LIHC project in the TCGA portal. Samples with a follow-up time less than 30 days or absent survival status were removed. GTF files used for annotation and thus distinction between lncRNAs and mRNAs were attained from the ENSEMBL database. 75 lymphangiogenesis-related genes with relevance scores greater than 1.5 were selected from the GeneCards database which was shown in Table 1. Selection of lymphangiogenesis-related genes was justified using Gene Ontology (GO) analysis.

2.2. Identification of lymphangiogenesis-related long noncoding ribonucleic acids pairs

Correlation analysis conducted between lymphangiogenesis-related genes and the entirety of lncRNAs was employed to screen lIrncRNAs. The standard for screening was defined as correlation coefficient greater than 0.4 and p-value less than 0.001. Subsequently, differential expression analysis was conducted among lIrncRNAs to identify differentially expressed lymphangiogenesis-related lncRNAs (DElncRNAs) with the criterion of FDR less than 0.05 and log2(|fold change (FC)|) greater than 2. To identify DElncRNAs pairs, DElncRNAs were cyclically paired in a lncRNA model able to predict prognosis and immunocheckpoint blockade of HCC(Xu et al., 2021). Further, signatures using two-biomarker combinations were proved to be more sensitive and applicable than the ones mentioned, partly because the latter required the specific expression of
Figure 1. Flow chart of the study.
value 0, would be yielded. The established zero-or-one matrix was subjected to further screening until DElncRNA pairs whose proportion of being value 1 or 0 was less than 20% or more than 80% were removed. The DElncRNA pairs removed are regarded as unnecessary for further analysis in that their expression ratios are considered the same in all samples (Hong et al., 2020).

Figure 2. DElncRNAs identification; Identifying differently expressed lymphangiogenesis-related IncRNAs (DElncRNAs) using the LIHC dataset from TCGA portal, as shown in the heatmap (A) and the volcano plot (B) (C) GO terms indicated the selected genes were lymphangiogenesis-related.
2.3. Construction and verification of the prognostic signature using differently expressed lymphangiogenesis-related long noncoding RNA pairs

Uni-Cox regression was performed for DElncRNA pairs left in the matrix with a p-value less than 0.01 to pick out the ones with prognostic significance. LASSO regression analysis was adopted to avert overfitting. Multivariate Cox regression was conducted later to establish the prognosis signature, which was used to calculate risk scores for clinical samples with the formula: \( \text{risk score} = \sum_{i=1}^{n} \frac{\text{val}(i)}{\text{coef}(i)} \), where \( n \) means the quantity of DElncRNA pairs within the prognostic signature and \( \text{val}(i) \) and \( \text{coef}(i) \) represent the value yielded in the matrix and the regression coefficient, respectively. To evaluate the prognosis signature, ROC curves were plotted with corresponding AUCs calculated and comparisons made between this signature and other clinical variables. Patients were allocated into high- or low-risk group as per the cut-off point generated from the Akaike information criterion (AIC) values of 1-year ROC curve. Survival differences between the two risk groups were compared using Kaplan-Meier method and logrank test. Subsequently, univariate and multivariate analyses were performed to discern independently prognostic predictors of HCC patients, and more importantly, whether the risk score serves as one. The relation between the prognosis signature and other clinicopathological features was investigated using chi-square tests. Differences in risk scores among subgroups with these clinical characteristics were shown as box plots via Wilcoxon signed-rank test. A nomogram model was constructed using the two independently prognostic predictors unanimously identified by the univariate and multivariate analyses and validated by calibration graphs of 1-/2-/3-year comparing the actual survival probability of HCC patients and the one yielded by the nomogram (Wan et al., 2017).

2.4. Biofunction and pathways exploration

A coexpression network between lymphangiogenesis-related lncRNAs in the signature and their interactive mRNA previously identified was built using Cytoscape (version 3.9.0). The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were conducted with R packages \texttt{org.Hs.eg.db} (version 3.13.0) and \texttt{GOplot} and visualized by \texttt{ggplot2}, under the criterion of \( p-/q\)-values \(< 0.05\).

2.5. Tumoral infiltration of immune cells analysis

The correlation between the risk score and the tumoral infiltration of immune cells estimated by well-received methods at present inclusive of XCELL, TIMER, QUANTISEQ, MCPCOUNTER, EPIC, CIBERSORT – ABS and CIBERSORT was investigated for later visualization in the form of a lollipop diagram via Spearman correlation test, whose significance threshold was specified as \( p\)-value less than 0.05. Wilcoxon signed-rank test was employed to analyze, and boxplots were plotted to display the infiltration differences of immune cells between the two risk groups.

2.6. Immune checkpoint genes expression

A better understanding of the expression differences in immunocheckpoint genes between the high-risk and low-risk groups, several immunocheckpoint genes were chosen for analysis, including CD47, CD276, LAG3, CTLA4, PDCD1 and HAVCR2; the results are visualized as violin plots by the ggpubr R package.
2.7. Chemosensitivity evaluation

To examine the prognosis signature under clinical applications, half-maximal inhibitory concentration (IC50), which indicates the concentration of drug needed to inhibit tumor cells by 50%, was calculated for five chemotherapeutic drugs reportedly useful for patients with hepatocellular carcinoma, including doxorubicin, gemcitabine, mitomycin C and sorafenib. Boxplots were drawn to display the contrasts in the IC50s between high-risk and low-risk groups computed by Wilcoxon signed-rank test via pRRophetic R package.

2.8. Cell lines and cell culture

Human normal liver cells (LO2) and human HCC cell lines (HepG2, SK-Hep-1 and Huh7) were purchased from the Cell Bank (Cell Institute, Sinica Academia Shanghai, Shanghai, China) and validated by short tandem repeat (STR) profiling. Cells were either cultured in RPMI 1640 (Gibco, USA) or Dulbecco’s Modified Eagle’s Media (DMEM, Biological Industries, Israel) supplemented with 10% fetal bovine serum (FBS, Gibco, USA) and 1% penicillin-streptomycin and incubated at 37°C in 5% CO2.

2.9. Verification of LncRNA pairs by quantitative real-time polymerase chain reaction

The EZ-press RNA Puriﬁcation Kit (EZBioscience, Shanghai, China) was used to extract total RNA from the human normal and tumor cells following the manufacturer’s instructions. The purity and concentration of RNA extracted were measured using IMPLEN N60 Touch (IMPLEN, Germany). HiScript II Q RT SuperMix for qPCR (+gDNA wiper) (Vazyme, Nanjing, China) was used for reverse transcription and ChamQ Universal SYBR qPCR Master Mix (Vazyme, Nanjing, China) for qRT-PCR reaction under the LightCycler96 System (Roche, Germany). LncRNA pair expression ratios were calculated using the relative 2−ΔΔCt method with internal control of β-actin. The differences in the ratios in liver normal and tumor cells were compared by t-tests and visualized in bar graphs drawn with GraphPad Prism (version 9.0). See Table 2 for the list of qRT-PCR primer sequences.

3. Results

3.1. Differently expressed lymphangiogenesis-related long noncoding RNAs

The flow chart of this study was shown as Figure 1. Transcriptomic profiles and clinical information of 377 LIHC patients were retrieved from TCGA portal. After removing the patients whose follow-up time was less than 30 days, 349 patients were included for further analysis. Lymphangiogenesis-related genes were attained from the GeneCards database and subjected to Gene Ontology analysis, in which GO terms justiﬁed the gene selection in this fashion, before and after being used for coexpression analysis to screen out lrlncRNAs (Figure 2C). Under the criterion of FDR <0.05 and log2 |fold change (FC)| >2, a total of 32 lrlncRNAs were found, all of which were upregulated (Figure 2A, B).
3.2. Establishment and validation of lymphangiogenesis-related long noncoding RNA pairs and prognosis signature

Using the zero-or-one matrix that singly and cyclically paired DElncRNAs, 346 valid DElncRNA pairs were constructed, among which 16 pairs were extracted after LASSO regression analysis following a univariate Cox regression analysis (Figure 3A, B). Subsequently, a Cox proportional hazard model was constructed stepwise with 9 of the 16 pairs; results were shown as Figure 3C, D. To elucidate the biological functions and pathways relevant to the established signature, lncRNAs within the signature and their coexpressed mRNAs were used to build the network (Figure 4A). The GO and KEGG analyses therefore performed indicated close functional connection of the signature with chemotaxis and endothelial cell proliferation (Figure 4B), and participation in Rap1, Ras, MAPK, PI3K–Akt, Calcium and AGE–RAGE signaling pathway (Figure 4C). Further, AUC values were calculated for the ROC curves of all the 9 DElncRNA pairs, the maximum of which was 0.825 and was used for calculation of the AIC value to yield the optimal cut-off point (Figure 4D). ROC curves of 1-/2-/3-year were plotted, and other clinical features were also compared to assess the optimality of the signature (Figure 4E, F). The AUC of 1-year ROC curve was much greater than that of other clinical characteristics, endorsing the clinical significance the signature possesses based on DElncRNA pairs (Figure 4F). Using the cut-off point 1.580, patients were allocated to different risk groups depending upon their risk scores (Figure 5A). It can be seen from the scatterplot and boxplot that patients in the low-risk group had a higher chance of survival (Figure 5B). The conclusion was also supported by the survival curve showing the survival probability of the high-risk patients was significantly lower than that of the low-risk (Figure 5C, p < 0.001).

Subgroup survival analyses performed to reduce bias rendered similar results (Figure 6). To explore the relation between the risk score and other clinical features, Chi-square tests were conducted among such subgroups of clinical features as gender, age, clinical stage, T/M/N stage and Grade. The results were displayed in a heatmap presenting the extent to which these clinical features were related to the signature. Of all, clinical stage, Grade and T stage were the ones significantly associated (Figure 7A). Wilcoxon signed-rank tests performed for the same purpose rendered similar results except for M stage (Figure 7B-H), which may be
Figure 6. Survival analysis for subgroups of gender (A, B), age (C, D), grade (E, F), stage (G, H), N stage (I, J), M stage (K, L) and T stage (M–O). P-values in all subgroups indicated statistical significance.
Figure 7. Association of the signature with clinicopathological features (A) The heatmap showed that grade, clinical stage and T stage were significantly related to the risk score (B–H). Boxplots using Wilcoxon signed-rank tests agreed that T stage (B), stage (C) and grade (D) significantly correlated with the risk score, while M stage (E), age (F) and gender (G) did not, with the exception of N stage (H). Univariate (I) and multivariate (J) Cox regression analyses discerning independently prognostic predictors.
attributed to insufficient cases of M1. To validate the signature without other clinical characteristics affecting the outcome, univariate (Figure 7I) and multivariate (Figure 7J) Cox regression analyses were adopted to discern independently prognostic predictors of LIHC patients and the results from both analyses agreed that risk score (Univariate: \( p < 0.001, \text{HR} = 1.208, \text{CI}(1.158–1.259) \); Multivariate: \( p < 0.001, \text{HR} = 1.190, \text{CI}(1.139–1.244) \)) and stage (Univariate: \( p < 0.001, \text{HR} = 1.808, \text{CI}(1.463–2.234) \); Multivariate: \( p < 0.001, \text{HR} = 1.652, \text{CI}(1.324–2.061) \)) were independent predictors of prognosis. Therefore, the risk score calculated using the signature works decently as recognized clinical predictors. To ameliorate the prediction of LIHC patients' survival, another independently prognostic predictor stage was combined with the risk score to form a nomogram model (Figure 8A). For instance, the estimated 3-year survival rate of a stage III LIHC patient with a risk score of 14 is less than 5%. Calibration graphs of 1-/2-/3-year comparing the actual survival probability of the patients and the one predicted by the nomogram revealed that differences between them were marginal (Figure 8B-D), suggesting the applicability of the nomogram model.

### 3.3. Exploration of relation between risk score and tumoral infiltration of immune cells

Given the involvement of IncRNA in the regulation of tumoral infiltration of immune cells, the relation between the IncRNA-based risk score and immune cells infiltration was explored using Wilcoxon signed-rank tests and Spearman correlation. A detailed list of immune cell types with significant Spearman correlation coefficients was presented in Figure 9A. It was revealed that high risk correlated with higher tumoral infiltration of macrophages (Figure 9B), Th2 cells (Figure 9C), myeloid dendritic cells (Figure 9D), Treg cells (Figure 9E) and neutrophils (Figure 9F), and with lower tumoral infiltration of CD8+ naïve cell (Figure 9G), hematopoietic stem cells (Figure 9H), endothelial cells (Figure 9I) and central memory T cells (Figure 9J).

### 3.4. Risk-related expression of immunocheckpoint genes

To find out if the risk score could be used to predict immune checkpoint blockage therapy, the expression of immunocheckpoint genes in the high-risk and low-risk groups were visualized comparatively in violin plots. It was shown that the expression of HAVCR2 (Figure 10A, \( p < 0.001 \)), CD47 (Figure 10B, \( p < 0.001 \)) and CD276 (Figure 10C, \( p < 0.001 \)) were significantly higher in high-risk group while the expression difference of the rest of the genes analyzed (Figure 10B-F, \( p > 0.05 \)) showed no statistical significance.

### 3.5. Chemotherapeutic prediction using risk score

IC50 values evaluating the chemosensitivity in LIHC patients were computed and the differences between the high-risk and low-risk groups were analyzed using Wilcoxon signed-rank test. As shown in the box-plots, the half inhibitory concentration of gemcitabine (Figure 11A, \( p < 0.001 \)), doxorubicin (Figure 11B, \( p < 0.01 \)), mitomycin C (Figure 11C, \( p < 0.001 \)) and sorafenib (Figure 11D, \( p < 0.001 \)) were significantly higher in high-risk group while the expression difference of the rest of the genes analyzed (Figure 10B-F, \( p > 0.05 \)) showed no statistical significance.

### 3.6. Validating expression-ratios of RNA pairs by quantitative real-time polymerase chain reaction

To validate the ratios of the IncRNA pairs expression, qRT-PCR was performed in normal and tumoral liver cell lines. As shown in Figure 12, the expression ratio of AC068506.1|LENG8-AS1 was significantly
Figure 9. Association of the signature with tumoral infiltration of immune cells (A) The correlation of the risk score with many types of tumor-infiltrating cells (B) High risk correlated with higher tumoral infiltration of macrophages (B), Th2 cells (C), myeloid dendritic cells (D), Treg cells (E) and neutrophils (F) and lower infiltration of CD8+ naïve cell (G), hematopoietic stem cells (H), endothelial cells (I) and central memory T cells (J).
elevated, while the ratio of AC006042.1| AL355488.1 was significantly decreased in the high-risk group, according with their hazard ratios in the signature, which indicated that the lncRNA pairs are worthy of further investigation.

4. Discussion

As a major cause of tumor-related mortality globally, HCC can be highly metastatic and recurrent, which restricts patients’ long-term survival (Vogel et al., 2019). Lymphangiogenesis, which describes the growth of new lymphatic vessels, has been shown to relate to metastases and unsatisfactory prognosis in a variety of human tumors, including melanoma, prostate and breast cancers (Rinderknecht and Detmar, 2008). More importantly, recent evidence has revealed that lymphangiogenesis facilitates metastasis in HCC (Yu et al., 2017). As lymphangiogenesis is rarely observed in healthy adults, therapies targeting lymph vessel formation should have the advantage of not intervening normal physiology (Mumprecht and Detmar, 2009). Anti-lymphangiogenic strategies have been developed over the decade to hinder lymphatic metastasis and currently proceed to the stage of clinical trials (Dieterich and Detmar, 2016). There is sufficient experimental evidence that drugs blocking the lymphangiogenic axis reduce tumor metastasis, lymphatically and distantly (Burton et al., 2008; Caunt et al., 2008). Together with the recent discovery that lymphangiogenesis regulates specific immune responses, it is tempting to develop lymphangiogenesis-related biomarkers as latent diagnostic and therapeutic targets for patients with HCC. In recent years, signatures based on specific expression of certain transcripts were proposed to predict the prognosis of malignancy, which required complicated calibration before clinical application (Wu et al., 2021; Xia et al., 2021; Xu et al., 2021). In the present study, we established a lymphangiogenesis prognostic signature for HCC patients by taking advantage of the relative expression of lncRNA pairs, thus allowing better practicability for clinical use.

First, differently expressed lrlncRNAs were sifted using correlation analysis conducted between lr-genes and the whole of lncRNAs with the data retrieved from TCGA portal. Next, valid lrlncRNA pairs were sifted using the zero-or-one matrix and the ones with prognostic significance were screened out and the prognostic signature was established after a series of computations. To test the signature, ROC curves of 1-/2-/3-year were plotted and compared to that of other clinical features such as gender, age and stage. The patients were allocated into high-risk or low-risk group as per the cut-off point computed using the AIC value. Subsequently, the relationship was investigated between the risk score and survival, clinical features, tumoral infiltration of immune cells, expression of immunocheckpoint inhibitor genes and chemosensitivity. Some of the differently expressed lrlncRNAs included in the signature have already been shown to be important players in HCC. Recent evidence indicated that LINC00205 promoted proliferation of HCC cells by targeting miR-122–5p or miR-26a-5p (Zhang et al., 2019; Cheng et al., 2021). LINC00665 was reported to increase malignancy of HCC through the activation of NF-κB signaling (Ding et al., 2020). LncRNA MYLK-AS1 was also found to facilitate tumor progression of HCC through miR-424–5p/E2F7 axis or EGFR/HER2-ERK1/2 signaling pathway (Liu et al., 2020; Teng et al., 2020). Moreover, expression ratios of AC068506.1| LENG8-AS1 and AC006042.1| AL355488.1 were validated by qRT-PCR in various HCC cell lines, which along with the above studies suggests that the lrlncRNAs within the established signature can be worthy of further investigation.

Tumor microenvironment embraces a broad spectrum of intricate interactions between immune cells, tumor cells and stroma, in which lymphangiogenesis also plays important roles in regulating antitumor immunity (Marin-Acevedo et al., 2018; Garnier et al., 2019). Immune cell
activation and infiltration in HCC affect response to anti-tumor blockade and relate to prognosis and therapeutic efficacy (Kurebayashi et al., 2018). To investigate the relation between tumoral infiltration of immune cells and risk score, the following acknowledged methods were used, including TIMER (Li et al., 2017), XCELL (Aran et al., 2017), QUANTISEQ (Finotello et al., 2019), CIBERSORT-A (Tamminga et al., 2020), CIBERSORT (Newman et al., 2015), MCPcounter (Dienstmann et al., 2019) and EPIC (van Veldhoven et al., 2011). In particular, the result revealed that high risk correlated with higher tumoral infiltration of macrophages, Th2 cells, myeloid dendritic cells, Treg cells and neutrophils and lower infiltration with CD8+ naïve cell, hematopoietic stem cells, endothelial cells and central memory T cells. A study recently demonstrated the possibility of predicting the therapeutic benefits of immunotherapy and chemotherapy using the immunogenomic analysis-derived immune scores (Dai et al., 2020). The lirncRNA signature proposed that high risk was linked with chemosensitivity to such therapeutics as cisplatin, doxorubicin, gemcitabine, mitomycin C and sorafenib, and with high expression of CD47, CD276 and HAVCR2.

Figure 11. Association of the signature with chemosensitivity; IC50 of gemcitabine (A), doxorubicin (B), mitomycin C (C) and sorafenib (D) were significantly lower in the high-risk group.

Figure 12. Verification of expression ratios in lymphangiogenesis-related lncRNA pairs; qRT-PCR results for the expression ratios of AC006042.1| AL355488.1 (A) and AC068506.1| LENG8-AS1 (B).
CD26 molecule, an immune checkpoint also known as B7–H3, inhibits anti-tumor immunity and promotes progression. In the tumor microenvironment, B7–H3 was capable of inhibiting Th1 activation and promoting Th2 differentiation (Feng et al., 2021). The shift from Th1 to Th2 was reported to promote cancer progression (Kumar et al., 2017). The effect of regulatory T cells on anti-tumor immune response suppression has been established (Tanaka and Sakaguchi, 2017). Dendritic cells exposed to regulatory T cells upregulate B7–H3 expression, producing an inhibitory phenotype (Mahnke et al., 2007). A plausible hypothesis was thus proposed that immunotherapy targeting B7–H3 allows the reversal of Th1/Th2 and reversal of DC inhibitory phenotype. The results in the study were in tune with previous evidence and together these findings suggest that the signature could potentialize clinical immunotherapy and chemotherapy guiding for patients with HCC.

However, we acknowledged as well some weaknesses and limitations in the present article. The raw data for initial analyses were relatively limited since it was merely retrieved from the TCGA portal due to the unavailability of datasets with simultaneous inclusion of lncRNA expression, clinicopathological features, and survival endpoints for HCC patients. Moreover, owing to the expression difference in samples, which might bring unsoundness to the ultimate model, external data validation was required, regardless of the zero-or-one matrix constructed to diminish sample errors sourcing and methods used for optimality testing. Subsequent molecular biological experiments are necessary and under consideration to further investigate the roles of DElncRNAs in HCC advancement and the underlying mechanisms. To improve the prognostic-predictive value of the signature, clinical cases would be preferred, the predictive value of the signature, clinical cases would be preferred, the

Subsequent molecular biological experiments are necessary and under consideration to further investigate the roles of DElncRNAs in HCC advancement and the underlying mechanisms. To improve the prognostic-predictive value of the signature, clinical cases would be preferred, the predictive value of the signature, clinical cases would be preferred, the

In conclusion, the present study demonstrated that a novel lncRNA-pairs based signature without test platforms limitations and not requiring specific expression levels of selected transcripts potentialize prognosis prediction of HCC and may help screen patients suitable for anti-tumor immunotherapy and chemotherapy.

Declarations

Author contribution statement
Jincheng Cao: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.
Yanni Xu: Performed the experiments; Analyzed and interpreted the data; Wrote the paper.
Xiaodi Liu: Conceived and designed the experiments; Analyzed and interpreted the data; Wrote the paper.
Yan Cai: Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.
Baoming Luo: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Funding statement
Prof Baoming Luo was supported by National Natural Science Foundation of China [82171944 & 81873899], Natural Science Foundation of Guangdong Province [2021A1515012611].

Data availability statement
Data associated with this study has been deposited at The Cancer Genome Atlas; Ensembl genome database; GeneCards database.

Declaration of interest’s statement
The authors declare no conflict of interest.
regulatory T cells: role of B7-H3 expression and antigen presentation. Eur. J. Immunol. 57 (8), 2117–2126.

Marín-Acevedo, J.A., Dholaria, B., Soyan, A.E., Knutson, K.L., Chamuji, S., Loh, Y., 2018. Next generation of immune checkpoint therapy in cancer: new developments and challenges. J. Hematol. Oncol. 11 (1), 39.

McGlynn, K.A., Petrick, J.L., El-Serag, H.B., 2021. Epidemiology of hepatocellular carcinoma. Hepatology (Baltimore, Md) 73 (Suppl 1), 4–13.

Ohkuma, T., Peters, S.A.E., Woodward, M., 2018. Sex differences in the association between diabetes and cancer: a systematic review and meta-analysis of 121 cohorts including 20 million individuals and one million events. Diabetologia 61 (10), 2140–2154.

Petrick, J.L., Campbell, P.T., Koshiol, J., Thistle, J.E., Andreotti, G., Beane-Freeman, L.E., et al., 2018. Tobacco, alcohol use and risk of hepatocellular carcinoma and intrahepatic cholangiocarcinoma: the Liver Cancer Pooling Project. Br. J. Cancer 118 (7), 1005–1012.

Qin, L.X., Tang, Z.Y., 2002. The prognostic molecular markers in hepatocellular carcinoma. World J. Gastroenterol. 8 (3), 385–392.

Rinderknecht, M., Detmar, M., 2008. Tumor lymphangiogenesis and melanoma metastasis. J. Cell Physiol. 216 (2), 347–354.

Statello, L., Guo, C.J., Chen, L.L., Huarte, M., 2021. Gene regulation by long non-coding RNAs and its biological functions. Nat. Rev. Mol. Cell Biol. 22 (2), 96–118.

Suzuki-Inoue, K., Tsukiji, N., Otake, S., 2020. Crosstalk between hemostasis and lymphangiogenesis. J. Thromb. Haemostasis 18 (4), 767–770.

Tammings, M., Hiltermann, T.J.N., Schuring, E., Timens, W., Fehrmann, R.S., Groen, H.J., 2020. Immune microenvironment composition in non-small cell lung cancer and its association with survival. Clin Transl Immunology 9 (6), e1142.

Tanaka, A., Sakaguchi, S., 2017. Regulatory T cells in cancer immunotherapy. J. Exp. Clin. Cancer Res. 36 (1), 235.

Teng, F., Zhang, J.X., Chang, Q.M., Wu, X.B., Tang, W.G., Wang, J.F., et al., 2020. LncRNA MYLK-AS1 facilitates tumor progression and angiogenesis by targeting miR-424-5p/ E2F7 axis and activating VEGFR-2 signaling pathway in hepatocellular carcinoma. J. Exp. Clin. Cancer Res. 39 (1), 235.

Thelen, A., Jonas, S., Benckert, C., Weichert, W., Schott, E., Bochtler, C., et al., 2009. Tumor-associated lymphangiogenesis correlates with prognosis after resection of human hepatocellular carcinoma. Ann. Surg Oncol. 16 (5), 1222–1229.

Torrecilla, S., Sia, D., Harrington, A.N., Zhang, Z., Cabellos, L., Cornella, H., et al., 2017. Trunk mutational events present minimal intra- and inter-tumoral heterogeneity in hepatocellular carcinoma. J. Hepatol. 67 (6), 1222–1231.

Vellinga, T.T., Kranenburg, O., Frenkel, N., Ubinik, L., Marvin, D., Govaert, K., et al., 2017. Lymphangiogenic gene expression is associated with lymph node recurrence and poor prognosis after partial hepatectomy for colorectal liver metastasis. Ann. Surg. 266 (5), 765–771.

Vogel, A., Cervantes, A., Chau, I., Daniele, B., Llovet, J.M., Meyer, T., et al., 2019. Hepatocellular carcinoma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. Ann. Oncol. 30 (5), 871–873.

Wan, G., Ge, F., Chen, J., Li, Y., Geng, M., Sun, L., et al., 2017. Nomogram prediction of individual prognosis of patients with hepatocellular carcinoma. BMC Cancer 17 (1), 91.

Wu, Q., Li, Q., Zhu, W., Zhang, X., Li, H., 2021. Identification of autophagy-related long non-coding RNA prognostic signature for breast cancer. J. Cell Mol. Med. 25 (8), 4088–4098.

Xia, P., Li, Q., Wu, G., Huang, Y., 2021. An immune-related lncRNA signature to predict survival in glioma patients. Cell. Mol. Neurobiol. 41 (8), 365–375.

Xu, Q., Wang, Y., Huang, W., 2021. Identification of immune-related lncRNA signature for predicting immune checkpoint blockade and prognosis in hepatocellular carcinoma. Int. Immunopharmac. 92, 107333.

Yu, S., Lv, H., Zhang, H., Jiang, Y., Hong, Y., Xia, R., et al., 2017. Heparanase-1-induced shedding of heparan sulfate from syndecan-1 in hepatocarcinoma cell facilitates lymphatic endothelial cell proliferation via VEGF-C/ERK pathway. Biochem. Biophys. Res. Commun. 485 (2), 432–439.

Zhang, L., Wang, Y., Sun, J., Ma, H., Guo, C., 2019. LINCO0205 promotes proliferation, migration and invasion of HCC cells by targeting miR-122-5p. Pathol. Res. Pract. 215 (9), 152515.

Zhao, J., Lawless, M.W., 2013. Long noncoding RNAs in tumor-associated lymphangiogenesis. Cancers 12 (11), 3290.

Zhang, L., Wang, Y., Sun, J., Ma, H., Guo, C., 2019. LINC00205 promotes proliferation, migration and invasion of HCC cells by targeting miR-122-5p. Pathol. Res. Pract. 215 (9), 152515.

Zhou, R., Liang, J., Tian, H., Chen, Q., Yang, C., Liu, C., 2021. Development of a ferroptosis-related lncRNA signature to predict the prognosis and immune landscape of bladder cancer. Dis. Markers 2021, 1031906.