


PRIMARY RESEARCH

Open Access



# PIWIL4 and SUPT5H combine to predict prognosis and immune landscape in intrahepatic cholangiocarcinoma

Wenbo Zou<sup>1,2†</sup>, Zizheng Wang<sup>2†</sup>, Xiuping Zhang<sup>2†</sup>, Shuai Xu<sup>2,4†</sup>, Fei Wang<sup>2</sup>, Lincheng Li<sup>1,2</sup>, Zhaoda Deng<sup>1,2</sup>, Jing Wang<sup>3</sup>, Ke Pan<sup>3</sup>, Xinlan Ge<sup>3</sup>, Chonghui Li<sup>3</sup>, Rong Liu<sup>2\*</sup> and Minggen Hu<sup>2\*</sup> 

## Abstract

**Background:** Intrahepatic cholangiocarcinoma (ICC) is a fatal primary liver cancer, and its long-term survival rate remains poor. RNA-binding proteins (RBPs) play an important role in critical cellular processes, failure of any one or more processes can lead to the development of multiple cancers. This study aimed to explore pivotal biomarkers and corresponding mechanisms to predict the prognosis of patients with ICC.

**Methods:** The transcriptomic and clinical information of patients were collected from The Cancer Genome Atlas and Gene Expression Omnibus databases. Bioinformatic methods were used to identify survival-related and differentially-expressed biomarkers. Quantitative real-time PCR (qRT-PCR) and immunohistochemistry were used to detect the expression levels of key biomarkers in independent real-world cohorts. Subsequently, a prognostic signature was constructed that effectively distinguished patients in the high- and low-risk groups. Independent prognosis analysis was used to verify the signature's independent predictive capabilities, and two nomograms were developed to predict survival.

**Results:** PIWIL4 and SUPT5H were identified and considered as pivotal biomarkers, and the same expression trends of upregulation in ICC were also validated via qRT-PCR and immunohistochemistry in the separate real-world sample cohorts. The prognostic signature showed good predictive capabilities according to the area under the curve. The correlation of the biomarkers with the tumour microenvironment suggested that the high riskScore was positively related to the enrichment of resting natural killer cells and activated memory CD4<sup>+</sup>T cells.

**Conclusion:** In the present study, we demonstrated that PIWIL4 and SUPT5H could be used as novel prognostic biomarkers to develop a prognostic signature. This study provides potential biomarkers of prognostic value for patients with intrahepatic cholangiocarcinoma.

**Keywords:** Intrahepatic cholangiocarcinoma, Prognostic biomarkers, Bioinformatics, Targeted therapy

## Background

Intrahepatic cholangiocarcinoma (ICC) is a fatal hepatobiliary malignancy derived from the bile ducts, and has a difficult diagnosis, poor prognosis, and very high chance of mortality [1, 2]. It accounts for approximately 10–20% of all bile duct malignancies [3]. To date, an increasing incidence of ICC has been reported in most of the world, while most patients with ICC have

\*Correspondence: liurong301@126.com; huminggen@301hospital.com.cn

<sup>†</sup>Wenbo Zou, Zizheng Wang, Xiuping Zhang and Shuai Xu authors contributed equally to this work.

<sup>2</sup> Faculty of Hepato-Pancreato-Biliary Surgery, The First Medical Center of Chinese People's Liberation Army (PLA) General Hospital, No.28 Fuxing Road, Haidian District, Beijing 100853, China

Full list of author information is available at the end of the article



© The Author(s) 2021. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

underlying liver diseases such as primary sclerotic cholangitis and hepatitis B, etc. [4, 5]. Surgery is considered the only treatment for cure at present, whereas most patients are not viable candidates for surgery at the time of diagnosis [6, 7]. Chemotherapy is another effective treatment, but most patients develop resistance eventually [8]. Therefore, finding novel and effective treatments to strengthen the long-term prognosis of ICC is necessary and urgent. Recently, immunotherapies and targeted therapies have gradually come to researchers' attention [9–12]. A variety of therapeutic targets have been identified and validated, and some vital biomarkers have been successfully transformed into clinical applications, such as FGFR, IDH1/2, PD1, PDL1 [11, 13–15]. However, there are still many undiscovered biomarkers, which play an important role in tumour progression; therefore, exploring additional prognostic biomarkers will help us better understand disease progression.

RNA-binding proteins (RBPs) play an important role in critical cellular processes, such as RNA splicing, modification, transport, localisation, stability, degradation, and translation [16]. These are essential for cell development, differentiation, and metabolism [17]. Thus, failure of any one or more of the above-mentioned processes can lead to the development of multiple diseases, including cancer [18–20]. Recent studies have demonstrated that RBP-mediated RNA modifications are critical for cancer progression via various molecular mechanisms, and that aberrant expression of RBPs in multiple cancer types affects the expression and function of oncogenes and anti-oncogenes [17, 21, 22].

Several studies have revealed that some crucial RBPs regulate tumorigenesis and cancer progression. For instance, the Musashi proteins, including Musashi-1 and Musashi-2, have been demonstrated as key prognostic biomarkers. Elevated Musashi protein expression characterises a variety of solid tumours and is associated with metastasis, lymph node invasion, and poor prognosis [23]. Increasing evidence suggests that the RBP HuR also plays a crucial role in disease progression by targeting the binding sites of oncogenes, or antioncogenes, and by regulating the corresponding cell processes [24, 25]. Similarly, elevated HuR expression is positively correlated with malignant biological behaviour in cancer. Notably, Toyota et al. demonstrated that high cytoplasmic HuR expression is closely associated with poor survival and the decreased efficacy of chemotherapy in patients with surgically resected cholangiocarcinoma [26]. Consequently, increasing evidence indicates that RBPs are extensively involved in tumour progression in ICC. At present, several RBPs and their underlying mechanisms have been revealed in the initiation and progression of

cancer, while more promising targets for ICC need be identified and verified.

In the present study, we identified and validated prognosis-related biomarkers PIWIL4 and SUPT5H using bioinformatic methods based on multiple public databases and as a result developed a risk classification model to predict the long-term survival of ICC patients. Notably, the biomarkers were further verified in ICC and normal tissues via laboratory experiments, which showed a similar expression trend to each other.

## Materials and methods

### Data collection

The RNA sequences regarding cholangiocarcinoma, specifically TCGA-CHOL, and clinicopathological information were obtained from The Cancer Genome Atlas (<https://portal.gdc.cancer.gov/>) and allocated into the training cohort. The GSE107943 validation cohort's transcriptome and clinicopathological data was collected from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>). The list of 1542 RBPs was obtained based on a previous published study [27]. The mRNA matrix was annotated based on gene transfer format files from Ensemble using the Perl language. Tumour tissues with the pathological diagnosis of ICC and adjacent normal tissues were prospectively collected from the Chinese PLA General Hospital. Written informed consent was obtained from all patients. After surgical resection, all the tissues were immediately collected and stored at  $-80^{\circ}\text{C}$ , awaiting further preparation for qRT-PCR analysis. The ICC tissue microarray for immunohistochemistry (IHC), which contained 155 ICCs and 5 adjacent normal tissues, was purchased from Shanghai Outdo Biotech Company and approved by the ethics committee of Shanghai Outdo Biotech Company (Shanghai, China).

### Extraction of RBP expression and differential expression analysis

Based on the obtained list of RBPs and mRNA expression profiles, we extracted the RBPs with available mRNA expression profiles from TCGA-CHOL and GSE107943 cohorts to construct a new matrix for subsequent analyses. Then, using the “limma” package [28], we screened the differentially expressed RBPs (DE-RBPs), with the thresholds set as false discovery rate (FDR)  $< 0.001$  and  $|\log_2 \text{fold-change (FC)}| > 0.5$ . Gene Ontology (GO) and the Kyoto Encyclopaedia of Genes and Genomes (KEGG) enrichment analyses were used to annotate the upregulated and downregulated DE-RBPs via the “clusterProfiler” and “GOplot” packages [29, 30]. A protein–protein interaction network was constructed from the DE-RBPs using the STRING online database (<https://string-db>).

**Table 1** Primers used for quantitative real-time PCR

GeneName	Direction	Sequences (5'–3')
h18S	Forward	AACCCGTTGAACCCATT
h18S	Reverse	CCATCCAATCGGTAGTAGCG
PIWIL4	Forward	CCAAGACTGGCAGCTATACCA
PIWIL4	Reverse	ACCGTCGAATGCTTTTGCTTT
SUPT5H	Forward	TGATCCCACGCATCGACTAC
SUPT5H	Reverse	TGGAGGCCGCTTAAACTTCTT

org/) with a medium confidence threshold of 0.4, and was then visualised using Cytoscape software (version 3.7.2).

**Survival-related RBPs screening and copy-number variation (CNV) analysis**

The univariate Cox regression analysis was implemented to find the survival-related DE-RBPs via the “survival” package, and the threshold was set to a p value < 0.05. Multivariate Cox regression analysis was used to determine the optimal survival-related DE-RBPs, and regression coefficients were noted for subsequent analysis. Kaplan–Meier (KM) curves were used to visualise the prognostic value of each DE-RBP in the TCGA cohort. To reveal the frequency of CNV of optimal DE-RBPs, we downloaded the CNV data of TCGA-CHOL from UCSC Xena (<http://xena.ucsc.edu/>), and revealed the RBPs location using the “RCircos” package [31].

**Quantitative real-time PCR**

Total RNA extraction, cDNA synthesis, and qRT-PCR were performed according to the manufacturer’s protocol and our previous article [32]. h18S rDNA was used

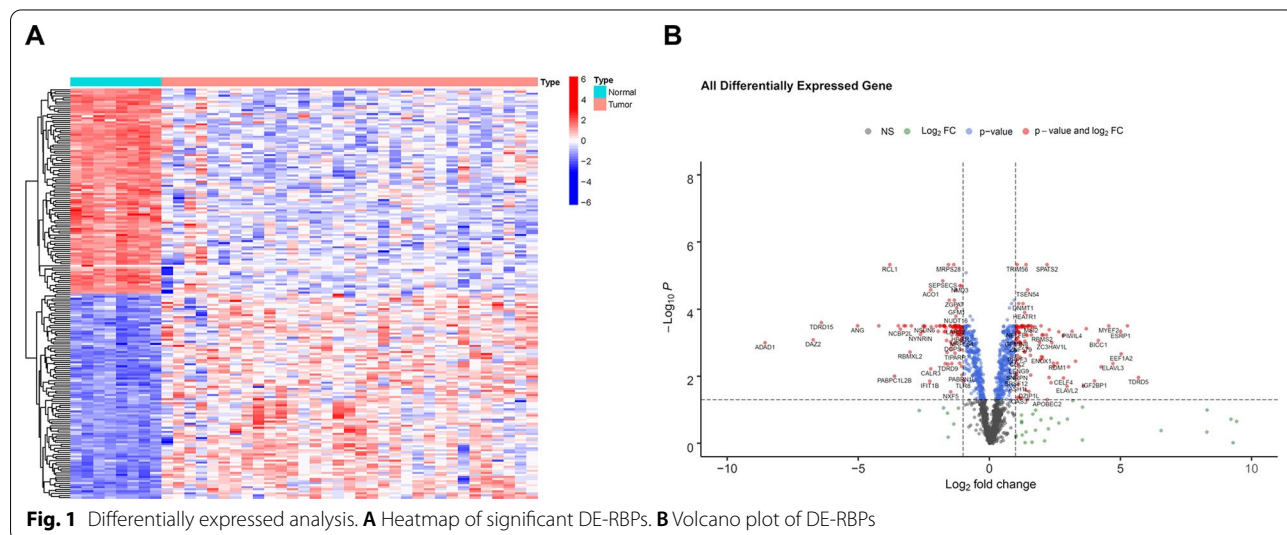
as an internal reference; finally, the cycle threshold (Ct) was recorded, and relative expression was calculated using the  $2^{-\Delta\Delta Ct}$  method. Primer sequences of PIWIL4, SUPT5H, and h18S rDNA are shown in Table 1.

**Immunohistochemistry**

All IHC staining for the tissue microarray was performed according to the manufacturer’s protocol. All images were obtained using a Leica Aperio XT digital pathology scanner (Leica, Wetzlar, Germany). The immunoreactive score (IRS) was calculated according to the ratio of positive cells and the staining intensity. This was used to evaluate expression levels between tumour and normal tissues. The ratio of positive cells was defined as follows: 0 (< 10%), 1 point (10–40%), 2 points (40–70%), and 3 (> 70%). The staining intensity was scored as 0 (negative), 1 (weakly positive), 2 (positive), and 3 (strongly positive). The two scores were added up to either 0–2 points to indicate weak expression, or 3–6 points, which was defined as strong expression. Detailed primary antibodies used for IHC are shown in Additional file 1: Table S1. The mean IRS from two random images of each sample was used to represent the final IRS using ImageJ software 1.53.

**Development and validation of RBP-related prognostic signature**

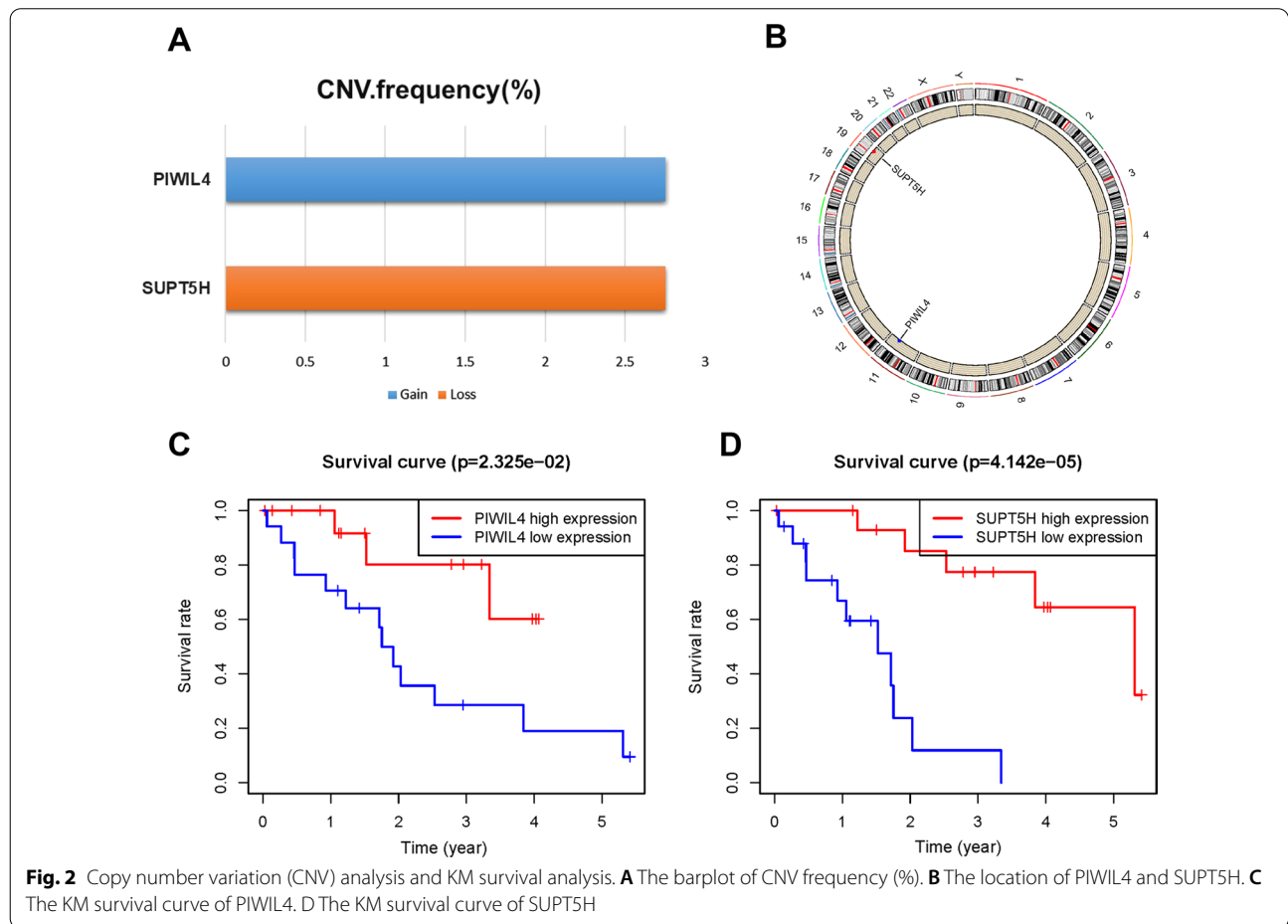
By combining the coefficients and expression levels of PIWIL4 and SUPT5H, we generated the riskScore of each patient using the following formula: riskScore =  $\sum_{i=1}^k \beta_i \text{exp}_i$ . All patients were grouped into high- or low-risk groups based on the median riskScore. The KM survival curve presented the predictive power, and



**Fig. 1** Differentially expressed analysis. **A** Heatmap of significant DE-RBPs. **B** Volcano plot of DE-RBPs

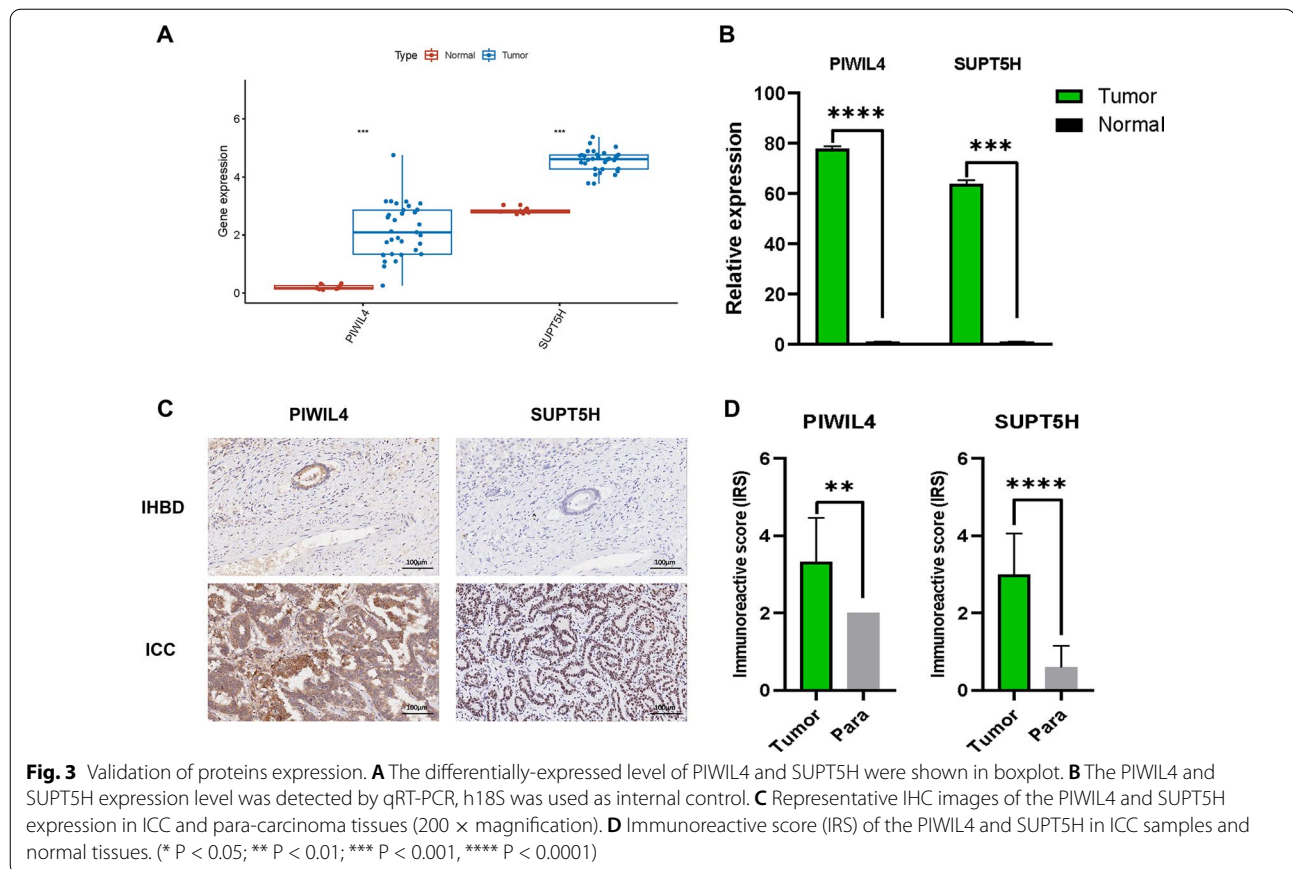
**Table 2** Univariate and multivariate Cox regression analysis for identifying independently prognostic biomarkers

Univariate Cox regression analysis					
Gene	HR	HR.95L	HR.95H	P value	
PIWIL4	0.4006	0.1856	0.8648	0.0198	
EIF4ENIF1	0.1630	0.0269	0.9895	0.0487	
SUPT5H	0.0046	0.0003	0.0674	< 0.001	
SCAF4	0.1656	0.0277	0.9894	0.0487	
Multivariate Cox regression analysis					
Gene	Coef	HR	HR.95L	HR.95H	P value
PIWIL4	- 0.6601	0.5168	0.2138	1.2492	0.1427
SUPT5H	- 4.8675	0.0077	0.0005	0.1178	< 0.001



the area under the curve (AUC) of the receiver operating characteristic (ROC) curve was used to verify the prediction accuracy using the “survivalROC” package. We calculated the AUC at 1, 2, and 3 years to verify the

model accuracy. Another independent gene expression profile, GSE107943, was used as the testing cohort for model validation. Finally, gene set enrichment analysis (GSEA) was performed to explore potential mechanisms,



and principal component analysis was used to efficiently downscale high-dimensional sequencing data.

#### Immune cell infiltration and chemotherapeutics efficacy analysis

We integrated acknowledged methods for evaluating the immune infiltration status in the TCGA-CHOL dataset, including TIMER [33], CIBERSORT [34], XCELL [35], QUANTISEQ [36], MCPcounter [37], and EPIC [38]. We analysed the correlation of the riskScore with immune cell infiltration [39], and used the TISIDB database (<http://cis.hku.hk/TISIDB/>) to explore the correlations between two RBPs expression and tumour immune microenvironment and immune-related markers.

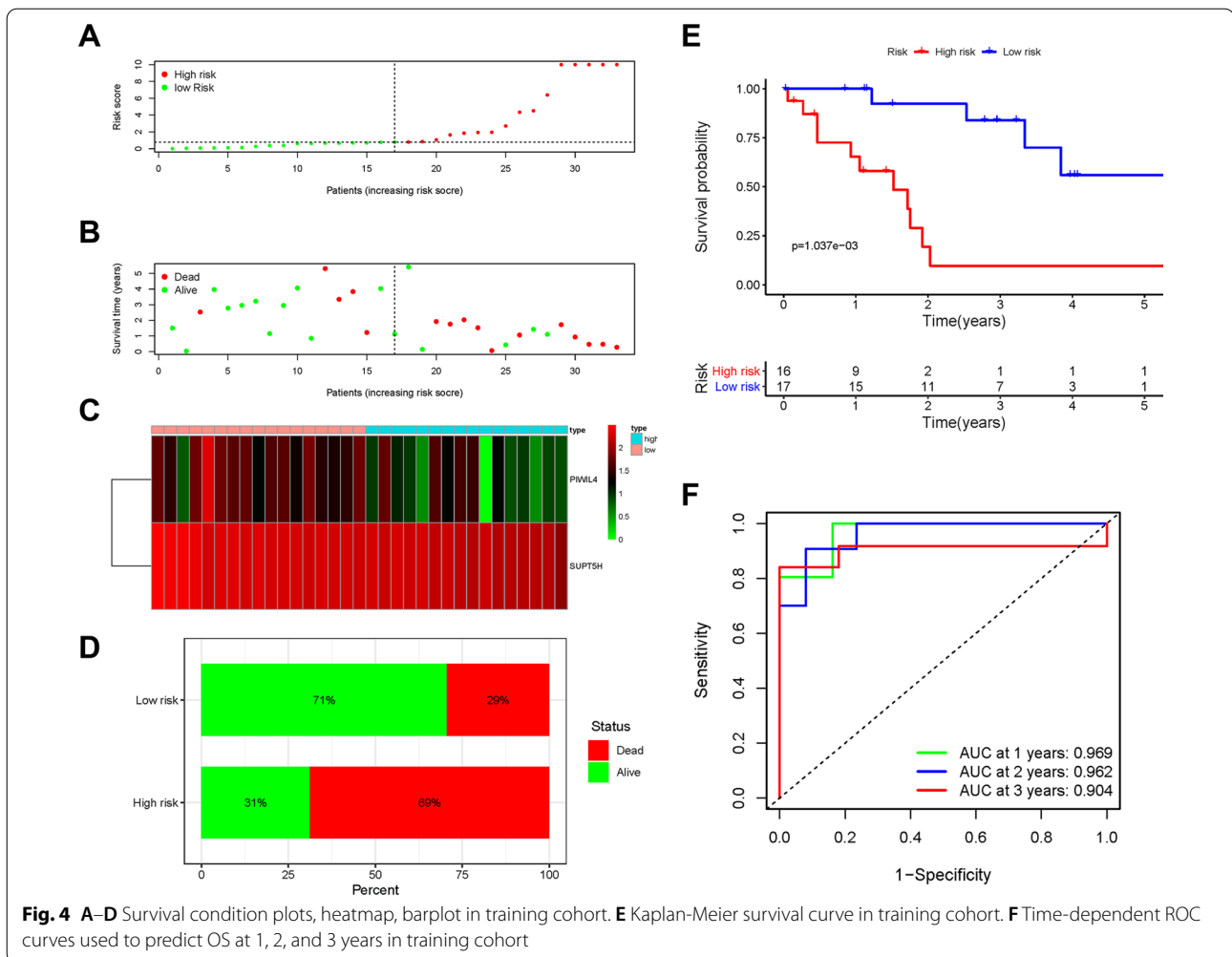
To evaluate the capability of the signature to predict chemotherapeutic efficacy in ICC, we compared the half-inhibitory concentration ( $IC_{50}$ ) difference of several chemotherapeutics in high- and low-risk groups. The Wilcoxon signed-rank test and the “ggplot2” and the “pRRophetic” packages were used to implement this process [40].

#### Optimal prognostic factors identification, clinical relevance analysis and nomogram construction

Univariate and multivariate regression analyses were performed to identify independent prognostic factors, including age, sex, tumour grade, American Joint Committee on Cancer (AJCC) stage, and peripheral nerve infiltration (PNI). Further survival analyses were also performed, stratified by clinical characteristics. Further, we generated two nomograms using the “rms” R package based on gene expression levels and clinicopathologic characteristics, respectively. The predictability of the nomograms was validated using the calibration curves.

#### Statistical analysis

Continuous variables were reported as medians (inter-quartile range) and were analysed by the Student’s t-test, while a log-rank test was used to perform the survival analysis. All statistical analyses and graphics were performed using R version 4.0.2 and its resource packages. Overall survival (OS) was defined as the length of time between surgery and death, or the last follow-up.



Two-sided  $p < 0.05$ , was considered statistically significant in all statistical tests.

## Results

### Data collection

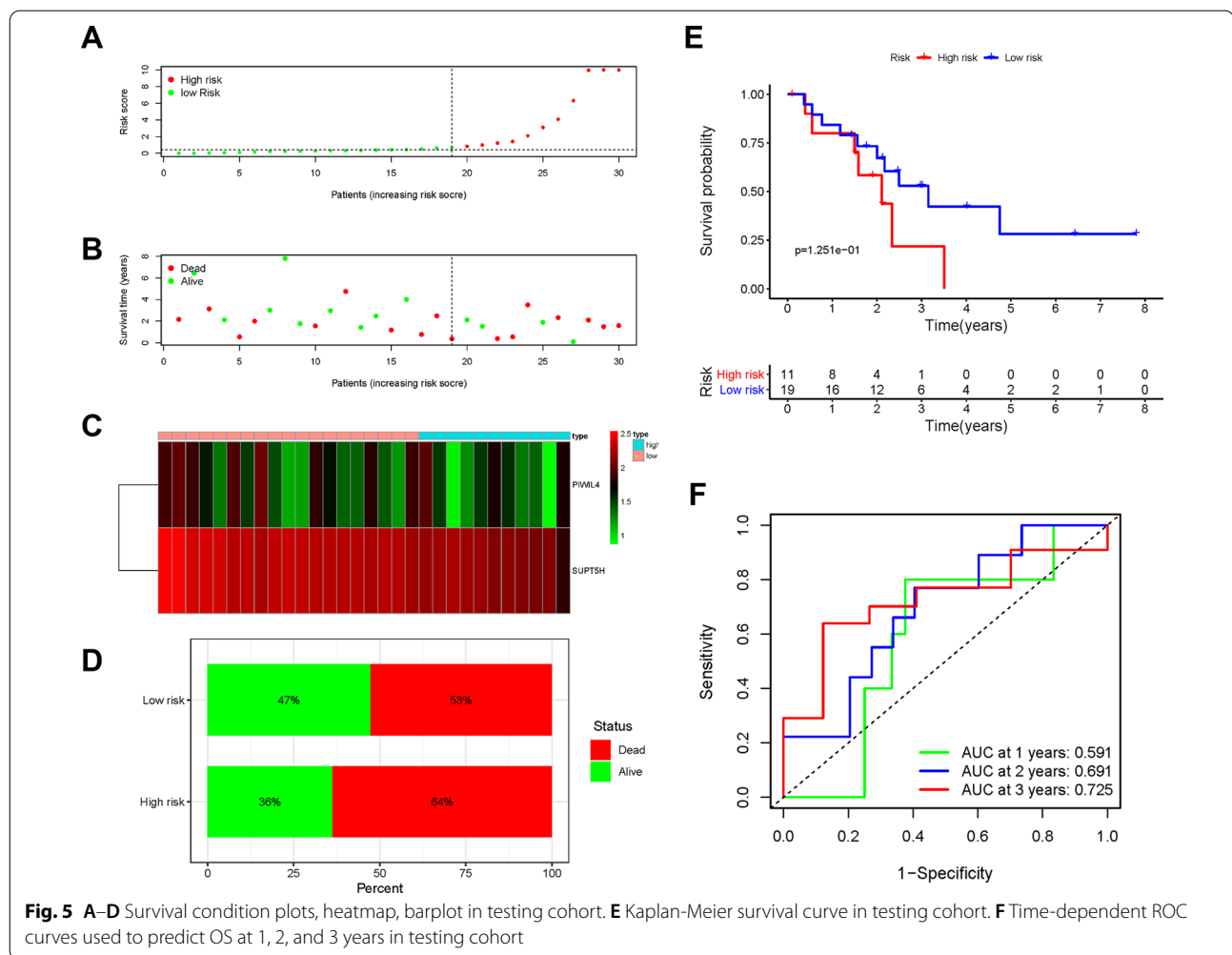
In total, 1523 RBPs with available mRNA expression profiles were extracted for subsequent analyses. The gene symbols of all obtained RBPs are shown in Additional file 1: Table S2. A total of 33 patient samples were obtained from TCGA and allocated into the training cohort, while an additional 30 patient samples were collected from GEO and grouped into the testing cohort. The detailed survival information (OS or last follow-up time) and the clinicopathologic characteristics of the patients involved in the two cohorts are shown in Additional file 1: Table S3.

### Differential expression analysis

Using the screening criteria, a total of 242 DE-RBPs were identified, of which 116 were upregulated and 126 were downregulated (Fig. 1A, B, Additional file 1: Table S4). As expected, GO and the KEGG pathway analyses demonstrated that DE-RBPs were correlated significantly with vital RNA regulatory processes, such as RNA transport and RNA degradation (Additional file 2: Fig. S1A–D). The PPI network and corresponding subgroup revealed associations between these DE-RBPs (Additional file 2: Fig. S2A–D, Additional file 1: Table S5).

### Survival-related RBPs screening and CNV analysis

Univariate and multivariate regression analyses identified two independently survival-related DE-RBPs, namely PIWIL4 and SUPT5H (Table 2), with the coefficients



being less than 0, suggesting that they were key protective factors in tumour progression. Then, we revealed the CNVs of PIWIL4 and SUPT5H, which are shown in Fig. 2A and Additional file 1: Table S6. Notably, we also presented their location in the genome (Fig. 2B), which could help us understand the roles they might play. Finally, the KM survival curve indicated that both PIWIL4 and SUPT5H were protective factors for survival classification ( $p < 0.05$ , Fig. 2C, D).

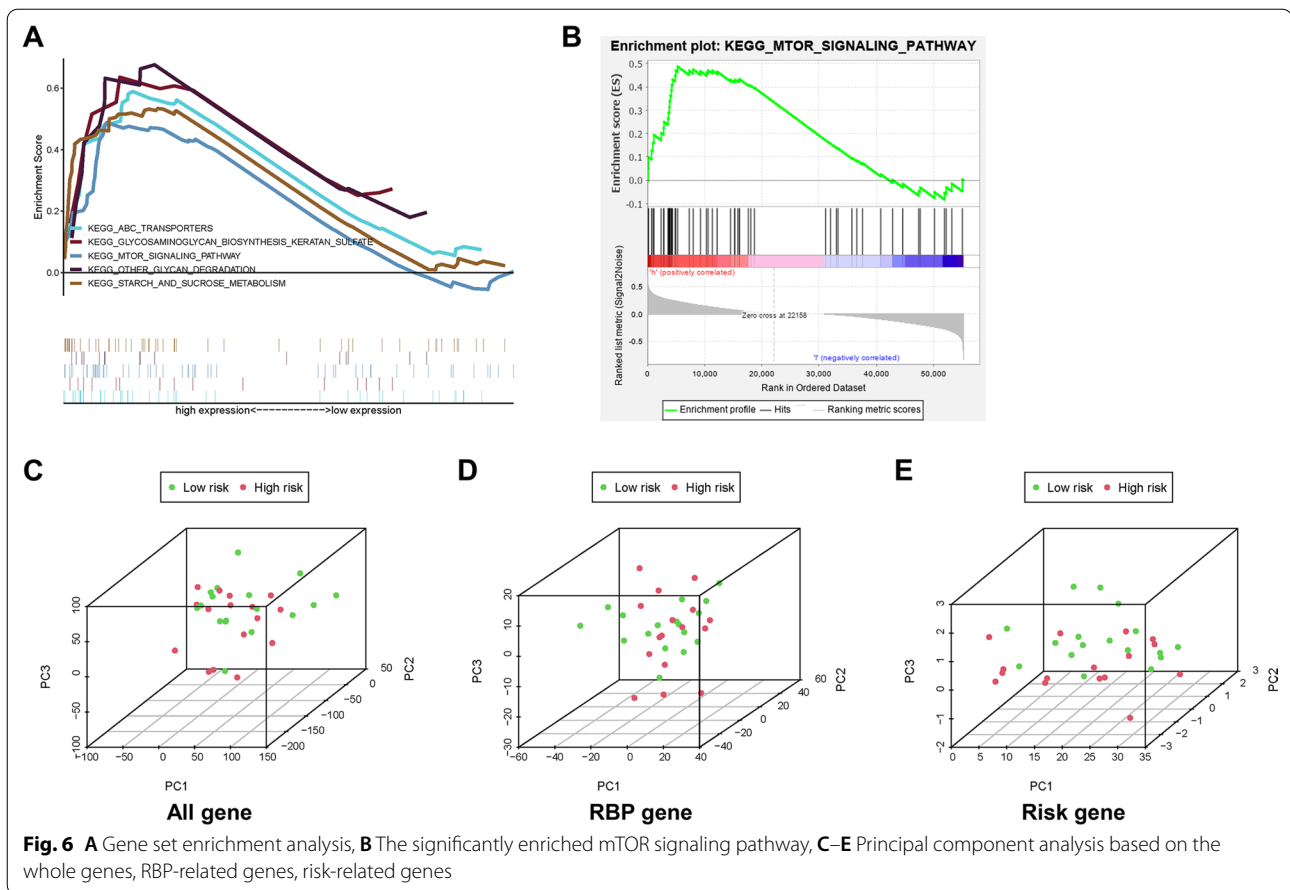
#### Validation of expression of PIWIL4 and SUPT5H in the independent cohorts

We verified the proteins PIWIL4 and SUPT5H using qRT-PCR and IHC in ICC. First, we analysed the transcriptomic data and demonstrated significantly upregulated levels of the proteins in ICC tissues (Fig. 3A), and the qRT-PCR results showed similar expression trends (Fig. 3B). We further performed IHC to validate the expression of these proteins. The IHC results showed

that PIWIL4 was located in the cytoplasm, and SUPT5H was located in the nucleus, while PIWIL4 and SUPT5H were both significantly upregulated in ICC (Fig. 3C, D).

#### Development and validation of RBP-related prognostic signature

An optimal prognostic signature was constructed. The riskScore was calculated as follows:  $\text{riskScore} = (-0.660) * \text{PIWIL4} + (-4.867) * \text{SUPT5H}$ . All patients in the training and testing cohorts were allocated into high- and low-risk groups using their median riskScore. In the training cohort, survival condition plots and bar plots showed the mortality difference among the risk groups and revealed the differential expression patterns of DE-RBPs (Fig. 4A–D). Moreover, the KM survival curve indicated that the high-risk group had a poorer prognosis and a shorter OS than the low-risk group ( $p < 0.01$ , Fig. 4E). The AUC calculated at 1, 2, and



3 years to verify the model accuracy were 0.969, 0.962, and 0.904, respectively (Fig. 4F).

In the testing cohort, survival condition plots and bar plots showed a similar difference in that the high-risk group had higher mortality than the low-risk group (Fig. 5A–D). The KM survival curve showed a survival difference between the two groups even if  $p > 0.05$  (Fig. 5E). The AUC at 1, 2, and 3 years were 0.591, 0.691, and 0.725, respectively (Fig. 5F). All the above results demonstrated that this signature could predict the prognosis of ICC patients with good predictive accuracy. In addition, these two RBPs were significantly correlated with one another in the training and testing cohorts (Additional file 2: Fig. S3).

#### Gene set enrichment analysis

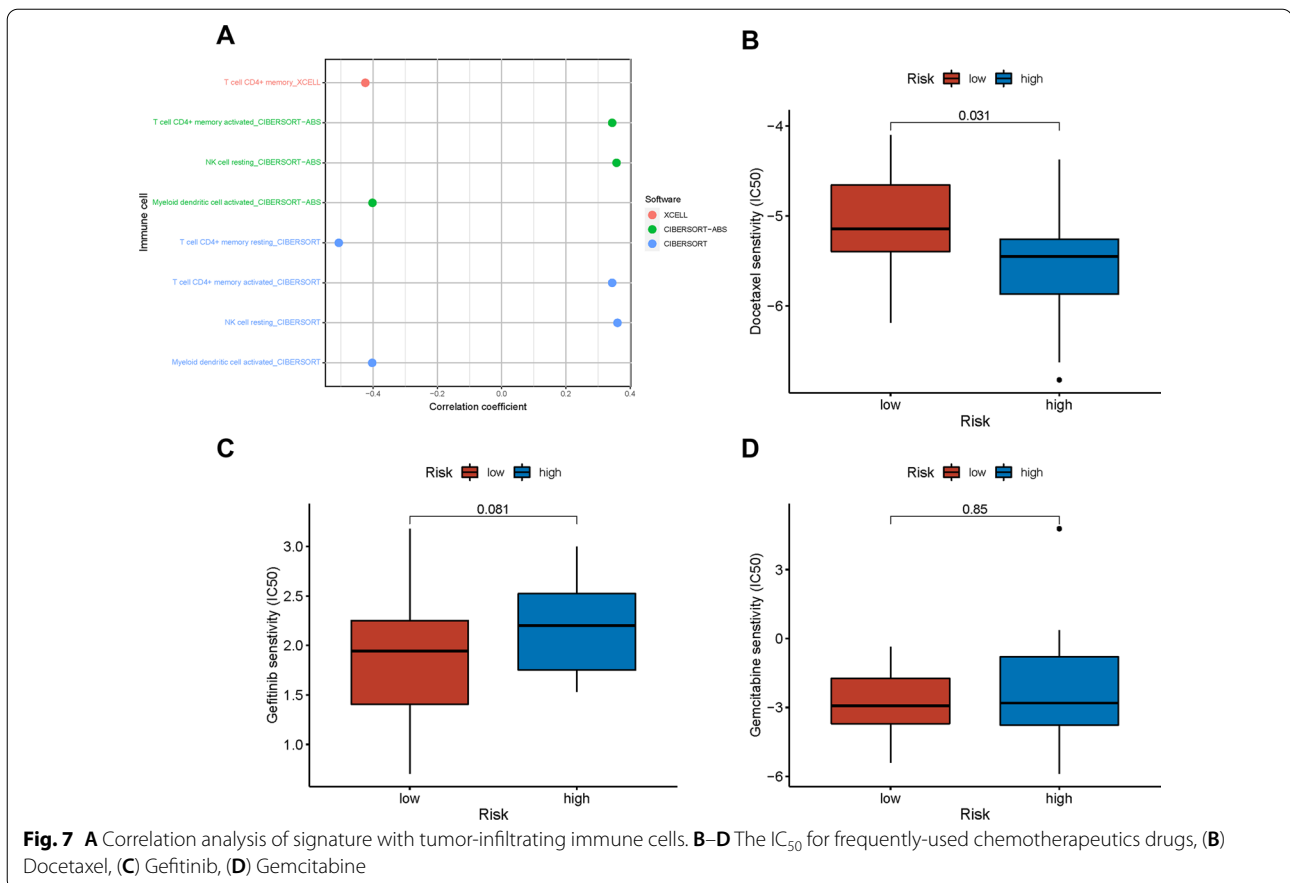
The significant enrichment pathways included ABC transporters, glycosaminoglycan biosynthesis keratan sulfate, mTOR signalling pathway, other glycan degradation, and starch and sucrose metabolism (Fig. 6A, Additional file 2: Fig. S4, and Additional file 1: Table S7). The mTOR signalling pathway has been reported to be the key pathway in tumorigenesis and progression (Fig. 6B).

The principal component analysis results showed that the model had a good risk classification (Fig. 6C–E).

#### Immune cell infiltration and chemotherapeutics efficacy analysis

We investigated the correlation of PIWIL4 and SUPT5H with the tumour microenvironment, and the results suggested that the high riskScore was positively related to the enrichment of resting natural killer (NK) cells and activated memory CD4+T cells. Meanwhile, the decreased enrichment of memory CD4+T cells activated myeloid dendritic cells and resting memory CD4+T cells (Fig. 7A). Notably, the immune correlation of the two DE-RBPs is shown in Table 3, and PIWIL4 was demonstrated to be differentially expressed in groups with or without immunotherapy for melanoma ( $p = 0.0385$ ) [41]. Subsequently, we aimed to investigate the relationship between the riskScore and common chemotherapeutic efficacy. The results showed that the low riskScore was positively correlated with the high  $IC_{50}$  of the chemotherapeutic agent docetaxel ( $p = 0.031$ ) and low  $IC_{50}$  of





gefitinib and gemcitabine (Fig. 7B–D); however the latter two agents were not statistically significant.

#### Clinical relevance analysis and nomogram construction

To further reveal the clinical predictive power of PIWIL4 and SUPT5H, univariate and multivariate Cox regression analyses were performed based on all the clinicopathologic characteristics and riskScore. The results suggested that the riskScore had an independent predictive value (Fig. 8A, B, Additional file 1: Table S8). Further survival analyses stratified by clinical characteristics were carried out, which showed that high-risk patients had poor long-term survival in all clinical characteristics for stratification survival analyses (Additional file 2: Fig. S5).

Next, we integrated the transcriptomic data and clinicopathologic characteristics into two nomograms for predicting 1, 2, 3 years survival rates (Fig. 8C, D). The multivariable ROC curve showed that the riskScore had the best predictive accuracy ( $AUC = 0.993$ , Additional file 2: Fig. S6A), and the calibration curves showed a good agreement between predicted and actual 1, 2, and 3 year survival rates (Additional file 2: Fig. 6B–G). The

clinical usefulness of the two nomograms was accurate and steady in predicting the long-term prognosis of patients with ICC.

#### Discussion

Similar to most malignancies, the long-term survival and therapeutic effects of patients with ICC remain poor [4]. Recently, with the increasing development of immunotherapy and targeted therapy, exploring novel tumour biomarkers has become a promising field of study [6]. In the present study, two prognostic biomarkers, PIWIL4 and SUPT5H, and their corresponding immune-mediated status for ICC were identified and validated. PIWIL4 and SUPT5H were both significantly upregulated in ICC and had the potential to become prognostic biomarkers. We then developed a prognostic signature, which could allocate the patients into high- and low-risk groups to distinguish long-term survival. In addition, the potential mechanism of the signature was revealed via GSEA, which showed that the biomarkers are involved in the mTOR signalling pathway in tumour progression. Increasing evidence has also demonstrated that the mTOR signalling pathway plays an essential role

**Table 3** Immune-related markers and mechanism analysis of PIWIL4 and SUPT5H

Gene	Types	Target	Spearman correlation analysis	p value	Survival	p value
PIWIL4	Chemokine receptor	CCL24	-0.434	p < 0.001		p = 0.00626
		CCR9	0.344	0.0408		
SUPT5H	Chemokine	CXCL5	-0.419	0.0116		p = 0.0234
		CXCL12	0.405	0.0148		
		CXCL16	-0.359	0.0321		
		IL10RB	0.44	0.00775		
	Immunoinhibitor	PVRL2	0.52	0.00133		
		Immunostimulator	CD276	0.357	0.0331	
			CXCL12	0.405	0.0148	
		ENTPD1	-0.364	0.0297		
		TMEM173	-0.334	0.0473		
		TNFSF4	-0.337	0.0447		
		TNFSF13	-0.382	0.0222		
		CHOL_MHC	HLA-DOB	-0.367	0.0282	
	CHOL_TIL_TEM	Tem CD4 cells	-0.349	0.0376		
	CHOL_TIL_Th2	Th2 cells				
	CHOL_TIL_Th17	Th17 cells	-0.421	0.0112		

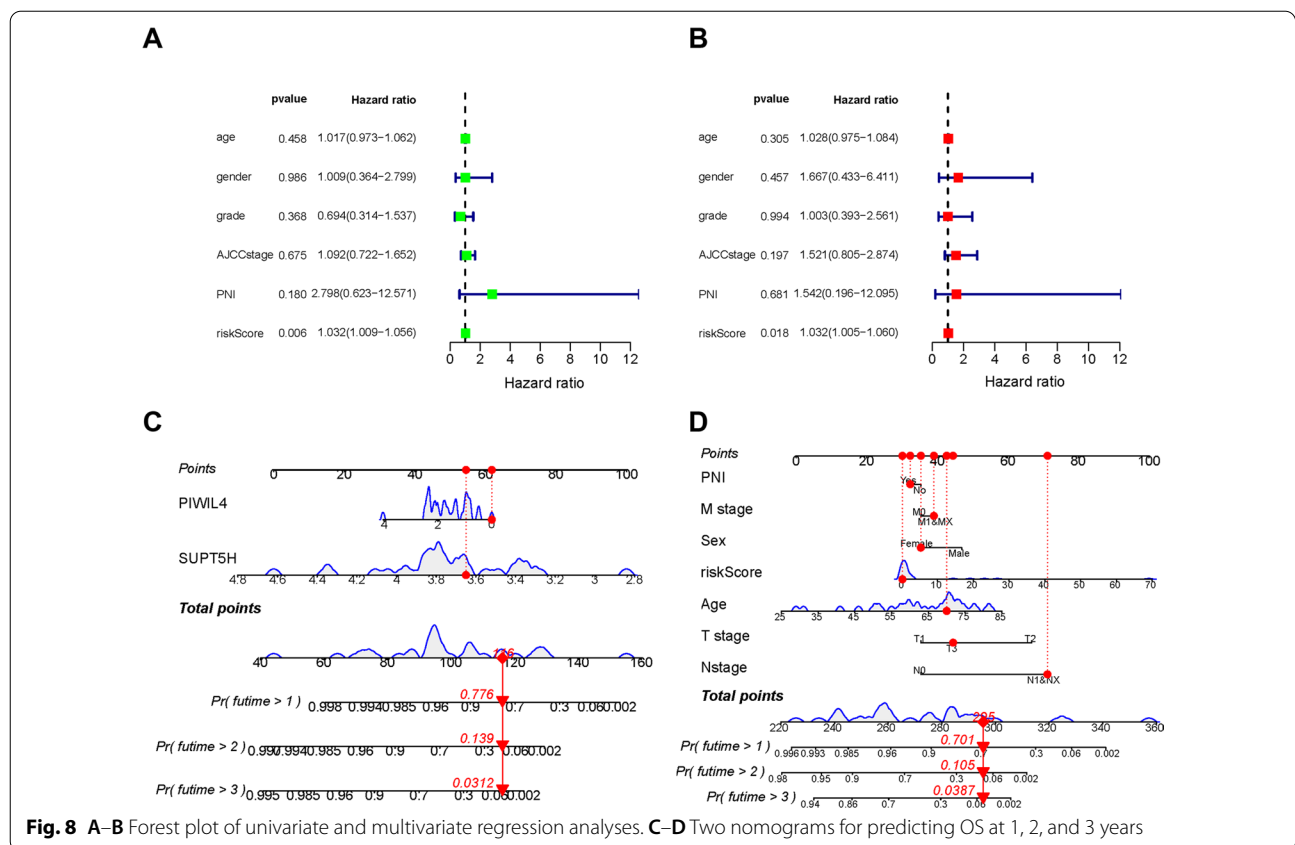
PIWIL4					
Cancer type	Drug	Group	Res vs N-Res	Log2(FC)	p value
Melanoma	Anti-PD-1 (pembrolizumab and nivolumab)	All	14 vs 12	-0.524	0.0385

Res The numbers of responders, NRes The numbers of non-responders

in the initiation and progression of ICC, and could be considered as a future therapy target [42, 43]. The prognostic signature was a powerful predictor of ICC prognosis, with its predictive capability validated in the testing cohort. Notably, we constructed two nomograms based on different detection methods, which provided different predictive tools for different patients. All calibration curves also showed an agreement between the probability of predictive survival and actual survival in 1, 2, and 3-year. Furthermore, the multivariate ROC curve showed that the riskScore had the best predictive accuracy.

PIWIL4 is a member of the P-element-induced wimpy testis (PIWI) protein family that was first discovered in germline cells [44]. PIWIL proteins bind a unique type of non-coding small RNA called piRNAs (PIWI-interacting RNAs) to form the piRNA/piwi complex, which influences protein regulation and epigenetic regulation, etc. [45]. Recently, PIWI family proteins have been considered as prognostic markers for various malignancies. A systematic review and meta-analysis showed that PIWI family proteins have the potential to indicate the prognosis of various cancer, and lower PIWIL4 expression levels indicate worse prognosis in cancer [46]. Iliev et al. demonstrated that decreased expression levels of PIWIL4 indicated worse long-term survival in patients with renal

cell carcinoma [47]. Li et al. also revealed that patients with low levels of PIWIL4 protein expression had a poor prognosis, and that PIWIL4 plays an important role in maintaining pancreatic cell homeostasis [48], which are results that are all consistent with our study. Notably, Mishra, N. K et al. study identified that PIWIL4 involve in DNA methylation and predict the prognosis, and can be used as a key prognostic biomarker in ICC [49]. In the present study, we explored the prognostic value of PIWIL4 based its RNA-binding protein mechanisms, and elucidated prognostic power of PIWIL4 from new perspective. thus, PIWIL4 should be considered and verified as an important biomarker in ICC. Recent evidence indicates that SUPT5H is a vital transcription promoter-binding protein involved in transcriptional elongation [50–52]. Lone et al. revealed a vital role of SUPT5H in regulating the expression levels of genes that control proliferation, migration, cell cycle, and apoptosis in breast cancer cases [53]. However, no study to the best of our knowledge has demonstrated the relationship between SUPT5H expression levels and the prognosis of patients with malignancies; our study being one of the first to identify SUPT5H as a prognostic biomarker in ICC, with low expression levels of SUPT5H reflecting poor prognosis, and it can be considered as a protective factor. This



**Fig. 8** A–B Forest plot of univariate and multivariate regression analyses. C–D Two nomograms for predicting OS at 1, 2, and 3 years

study provides novel research targets in tumour progression and prognosis of ICC.

To date, several prognostic signatures have been developed for predicting the survival of patients with ICC. In a previous study, Guo et al. comprehensively analysed and identified prognostic signatures, including seven mRNAs, for predicting recurrence in cholangiocarcinoma [54]. Mishra et al. identified nine genes which could also be strongly considered as prognostic markers of cholangiocarcinoma [49]. Huang et al. also identified three genes as pivotal tumour antigens of cholangiocarcinoma, which could benefit mRNA vaccine development [15]. Furthermore, our team and Xie et al. developed and validated a prognostic signature based on lncRNA-seq data [32, 55].

However, even though RBPs have been identified to have a crucial effects [20, 23, 56], a vital RBP-related risk model for ICC has not been reported yet. To our knowledge, we are the first to identify and construct the risk signature in ICC patients using two key RBPs. Notably, in the present study, this signature was verified in independent real-world cohorts.

In addition, the tumour immune microenvironment plays a pivotal role in the response to tumour immunotherapy [57]. Recently, a variety of studies has reported that immune cell infiltration and expression of immune

checkpoints directly affect patient prognosis [58–60]. Therefore, we further investigated the relationship between immune cell infiltration and the risk signature. The results suggested that the high riskScore had more enrichments of resting NK cells and activated memory CD4+ T cells. We also revealed the relationship between the immune landscape and each RBP. Therefore, this signature was able to depict the immune landscape in ICC. This signature could also predict the efficacy of chemotherapeutic agents like docetaxel. Consequently, the signature could be considered as an indicator of ICC patients who may benefit from immunotherapy and some chemotherapeutic drugs.

Although these findings and our model had a good capacity to predict the long-term survival of ICC patients, there are some limitations to this study. First, all transcriptomic data was obtained from a public database, and retrospective analysis was performed. Thus, selection bias was inevitable. Notably, we validated that PIWIL4 and SUPT5H expression pattern in ICC using the laboratory experiments, which indicated that both PIWIL4 and SUPT5H might involve tumour progression of ICC. Furthermore, we tentatively explored whether PIWIL4 and SUPT5H were highly related to the mTOR signaling pathway; however, the transcriptional regulation

mechanism of PIWIL4 and SUPT5H still needs to be revealed, as well as their effect on tumour progression and long-term survival needs to be clarified. Thus, further laboratory experiments for validation of the mechanism are required.

## Conclusion

PIWIL4 and SUPT5H were identified and validated as novel prognostic biomarkers via bioinformatics and laboratory experiments, and a signature was developed to predict the prognosis and risk classifiers of ICC patients. We also identified the immune landscape of the two markers in the ICC immune microenvironment. This study offers a promising perspective for exploring biomarkers in ICC.

## Abbreviations

ICC: Intrahepatic cholangiocarcinoma; RBPs: RNA binding proteins; TCGA: The Cancer Genome Atlas; GEO: Gene Expression Omnibus; AUC: Area under the curve; CNV: Copy number variation; Ct: Cycle threshold; FC: Fold change; GTF: Gene transfer format; GO: Gene ontology; GSEA: Gene set enrichment analysis; IHC: Immunohistochemistry; IRS: Immunoreactive score; ICI: Immune cell infiltration; KEGG: Kyoto Encyclopaedia of Genes and Genomes; MSI1: Musashi-1; MSI2: Musashi-2; OS: Overall survival; PCA: Principal component analysis; PNI: Peripheral nerve infiltration; ROC: Receiver operating characteristic; DE-RBPs: Differentially expressed RNA binding proteins; KM: Kaplan–Meier; qRT-PCR: Quantitative real-time PCR; NK: Natural killer cell; IC<sub>50</sub>: Half inhibitory concentration; PPI: Protein–protein interaction.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12935-021-02310-2>.

**Additional file 1: Table S1.** Primary antibodies used for immunohistochemistry (IHC). **Table S2.** Identification of RNA-binding proteins. **Table S3.** Clinical information of all patients with intrahepatic cholangiocarcinoma. **Table S4.** Differentially expressed analysis of RBP. **Table S5.** Detailed information protein–protein interaction network. **Table S6.** Copy number variation data obtained from TCGA database. **Table S7.** Gene set enrichment analysis. **Table S8.** Univariate and multivariate Cox regression analysis for identifying independently prognostic factors.

**Additional file 2: Figure S1.** Functional enrichment analysis of up and down RNA-binding proteins. **Figure S2. A** The protein–protein interaction (PPI) network; **B–D** subset visualization on Cytoscape. **Figure S3.** Spearman correlation analysis of PIWIL4 and SUPT5H in the TCGA and GEO databases. **Figure S4.** Gene set enrichment analysis. **Figure S5.** Kaplan–Meier survival analysis of the signature stratified by clinical characteristics. **Figure S6.** A Multivariable AUC values from ROC.

## Acknowledgements

We acknowledge TCGA, GEO, TISIDB and UCSC Xena databases for providing their platforms and contributors for uploading their meaningful datasets, and thank all individuals who participated in this study and donated samples.

## Authors' contributions

Study concept and design: H-MG, LR, Z-WB. Drafting of the manuscript: H-MG, LR, Z-WB. Acquisition of data, analysis, and interpretation of data: W-ZZ, Z-XP, XS, WF, L-LC, D-ZD, WJ, L-CH. Critical revision of the manuscript: H-MG, LR. Statistical analysis: G-XL, PK. Study supervision: HMG, LR. All authors read and approved the final manuscript.

## Funding

This study was supported by funds from National Key Sci-Tech Special Project of China (2018ZX10302204003).

## Availability of data and materials

The public data used to support the results of this study can be obtained from The Cancer Genome Atlas (TCGA) (<https://cancergenome.nih.gov/>), GEO database (<https://www.ncbi.nlm.nih.gov/geo/>), UCSC Xena (<http://xena.ucsc.edu/>) and TISIDB database (<http://cis.hku.hk/TISIDB/>).

## Declarations

### Ethics approval and consent to participate

This study was approved by the ethics committee of the Chinese PLA general Hospital and Shanghai Outdo Biotech Company. All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards. Written informed consent was obtained from all patients.

### Consent for publication

Consent for publication was obtained from all authors.

### Competing interests

The authors declare that they have no competing interests.

### Author details

<sup>1</sup>Medical School of Chinese PLA, Beijing, China. <sup>2</sup>Faculty of Hepato-Pancreato-Biliary Surgery, The First Medical Center of Chinese People's Liberation Army (PLA) General Hospital, No.28 Fuxing Road, Haidian District, Beijing 100853, China. <sup>3</sup>Institute of Hepatobiliary Surgery of Chinese PLA, Key Laboratory of Digital Hepatobiliary Surgery, PLA, Beijing, China. <sup>4</sup>School of Medicine, Nankai University, Tianjin, China.

Received: 23 July 2021 Accepted: 1 November 2021

Published online: 07 December 2021

## References

- Banales JM, Marin JJG, Lamarca A, Rodrigues PM, Khan SA, Roberts LR, Cardinale V, Carpino G, Andersen JB, Braconi C, et al. Cholangiocarcinoma 2020: the next horizon in mechanisms and management. *Nat Rev Gastroenterol Hepatol.* 2020;17(9):557–88.
- Nakeeb A, Pitt HA, Sohn TA, Coleman J, Abrams RA, Piantadosi S, Hruban RH, Lillemoe KD, Yeo CJ, Cameron JL. Cholangiocarcinoma. A spectrum of intrahepatic, perihilar, and distal tumors. *Ann Surg.* 1996;224(4):463–73.
- DeOliveira ML, Cunningham SC, Cameron JL, Kamangar F, Winter JM, Lillemoe KD, Choti MA, Yeo CJ, Schulick RD. Cholangiocarcinoma: thirty-one-year experience with 564 patients at a single institution. *Ann Surg.* 2007;245(5):755–62.
- Razumilava N, Gores GJ. Cholangiocarcinoma. *Lancet.* 2014;383(9935):2168–79.
- Karlsen TH, Folseraas T, Thorburn D, Vesterhus M. Primary sclerosing cholangitis - a comprehensive review. *J Hepatol.* 2017;67(6):1298–323.
- Rizvi S, Khan SA, Hallemeier CL, Kelley RK, Gores GJ. Cholangiocarcinoma - evolving concepts and therapeutic strategies. *Nat Rev Clin Oncol.* 2018;15(2):95–111.
- Rizvi S, Gores GJ. Pathogenesis, diagnosis, and management of cholangiocarcinoma. *Gastroenterology.* 2013;145(6):1215–29.
- Wei L, Wang X, Lv L, Liu J, Xing H, Song Y, Xie M, Lei T, Zhang N, Yang M. The emerging role of microRNAs and long noncoding RNAs in drug resistance of hepatocellular carcinoma. *Mol Cancer.* 2019;18(1):147.
- Klein O, Kee D, Nagrial A, Markman B, Underhill C, Michael M, Jackett L, Lum C, Behren A, Palmer J, et al. Evaluation of combination nivolumab and ipilimumab immunotherapy in patients with advanced biliary tract cancers: subgroup analysis of a phase 2 nonrandomized clinical trial. *JAMA Oncol.* 2020;6(9):1405–9.

10. Sirica AE, Gores GJ, Groopman JD, Selaru FM, Strazzabosco M, Wei Wang X, Zhu AX. Intrahepatic cholangiocarcinoma: continuing challenges and translational advances. *Hepatology*. 2019;69(4):1803–15.
11. Abou-Alfa GK, Macarulla T, Javle MM, Kelley RK, Lubner SJ, Adeva J, Cleary JM, Catenacci DV, Borad MJ, Bridgewater J, et al. Ivosidenib in IDH1-mutant, chemotherapy-refractory cholangiocarcinoma (ClarIDHy): a multicentre, randomised, double-blind, placebo-controlled, phase 3 study. *Lancet Oncol*. 2020;21(6):796–807.
12. Abou-Alfa GK, Sahai V, Hollebecque A, Vaccaro G, Melisi D, Al-Rajabi R, Paulson AS, Borad MJ, Gallinson D, Murphy AG, et al. Pemigatinib for previously treated, locally advanced or metastatic cholangiocarcinoma: a multicentre, open-label, phase 2 study. *Lancet Oncol*. 2020;21(5):671–84.
13. Rizvi S, Gores GJ. Emerging molecular therapeutic targets for cholangiocarcinoma. *J Hepatol*. 2017;67(3):632–44.
14. Loeuillard E, Yang J, Buckarma E, Wang J, Liu Y, Conboy C, Pavelko KD, Li Y, O'Brien D, Wang C, et al. Targeting tumor-associated macrophages and granulocytic myeloid-derived suppressor cells augments PD-1 blockade in cholangiocarcinoma. *J Clin Invest*. 2020;130(10):5380–96.
15. Huang X, Tang T, Zhang G, Liang T. Identification of tumor antigens and immune subtypes of cholangiocarcinoma for mRNA vaccine development. *Mol Cancer*. 2021;20(1):50.
16. Wang ZL, Li B, Luo YX, Lin Q, Liu SR, Zhang XQ, Zhou H, Yang JH, Qu LH. Comprehensive Genomic Characterization of RNA-Binding Proteins across Human Cancers. *Cell Rep*. 2018;22(1):286–98.
17. Qin H, Ni H, Liu Y, Yuan Y, Xi T, Li X, Zheng L. RNA-binding proteins in tumor progression. *J Hematol Oncol*. 2020;13(1):90.
18. Hodson DJ, Screen M, Turner M. RNA-binding proteins in hematopoiesis and hematological malignancy. *Blood*. 2019;133(22):2365–73.
19. de Bruin RG, Rabelink TJ, van Zonneveld AJ, van der Veer EP. Emerging roles for RNA-binding proteins as effectors and regulators of cardiovascular disease. *Eur Heart J*. 2017;38(18):1380–8.
20. Pereira B, Billaud M, Almeida R. RNA-binding proteins in cancer: old players and new actors. *Trends Cancer*. 2017;3(7):506–28.
21. Iino K, Mitobe Y, Ikeda K, Takayama KI, Suzuki T, Kawabata H, Suzuki Y, Horie-Inoue K, Inoue S. RNA-binding protein NONO promotes breast cancer proliferation by post-transcriptional regulation of SKP2 and E2F8. *Cancer Sci*. 2020;111(1):148–59.
22. Lujan DA, Ochoa JL, Hartley RS. Cold-inducible RNA binding protein in cancer and inflammation. *Wiley Interdiscip Rev RNA*. 2018. <https://doi.org/10.1002/wrna.1462>.
23. Kudinov AE, Karanicolis J, Golemis EA, Bumber Y. Musashi RNA-binding proteins as cancer drivers and novel therapeutic targets. *Clin Cancer Res*. 2017;23(9):2143–53.
24. Wang J, Guo Y, Chu H, Guan Y, Bi J, Wang B. Multiple functions of the RNA-binding protein HuR in cancer progression, treatment responses and prognosis. *Int J Mol Sci*. 2013;14(5):10015–41.
25. Xie M, Yu T, Jing X, Ma L, Fan Y, Yang F, Ma P, Jiang H, Wu X, Shu Y, et al. Exosomal circSHKBP1 promotes gastric cancer progression via regulating the miR-582-3p/HUR/VEGF axis and suppressing HSP90 degradation. *Mol Cancer*. 2020;19(1):112.
26. Toyota K, Murakami Y, Kondo N, Uemura K, Nakagawa N, Takahashi S, Sueda T. Cytoplasmic Hu-Antigen R (HuR) expression is associated with poor survival in patients with surgically resected cholangiocarcinoma treated with adjuvant gemcitabine-based chemotherapy. *Ann Surg Oncol*. 2018;25(5):1202–10.
27. Gerstberger S, Hafner M, Tuschl T. A census of human RNA-binding proteins. *Nat Rev Genet*. 2014;15(12):829–45.
28. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43(7):e47.
29. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS*. 2012;16(5):284–7.
30. Walter W, Sánchez-Cabo F, Ricote M. GPlot: an R package for visually combining expression data with functional analysis. *Bioinformatics*. 2015;31(17):2912–4.
31. Zhang H, Meltzer P, Davis S. RCircos: an R package for Circos 2D track plots. *BMC Bioinformatics*. 2013;14:244.
32. Zou W, Wang Z, Wang F, Li L, Liu R, Hu M. A metabolism-related 4-lncRNA prognostic signature and corresponding mechanisms in intrahepatic cholangiocarcinoma. *BMC Cancer*. 2021;21(1):608.
33. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, Li B, Liu XS. TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res*. 2020;48(W1):W509–14.
34. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA. Profiling tumor infiltrating immune cells with CIBERSORT. *Methods Mol Biol*. 2018;1711:243–59.
35. Aran D. Cell-type enrichment analysis of bulk transcriptomes using xCell. *Methods Mol Biol*. 2020;2120:263–76.
36. Finotello F, Mayer C, Plattner C, Laschober G, Rieder D, Hackl H, Krogsdam A, Loncova Z, Posch W, Wilflingseder D, et al. Molecular and pharmacological modulators of the tumor immune contexture revealed by deconvolution of RNA-seq data. *Genome Med*. 2019;11(1):34.
37. Dienstmann R, Villacampa G, Sveen A, Mason MJ, Niedzwiecki D, Nesbakken A, Moreno V, Warren RS, Lothe RA, Guinney J. Relative contribution of clinicopathological variables, genomic markers, transcriptomic subtyping and microenvironment features for outcome prediction in stage II/III colorectal cancer. *Ann Oncol*. 2019;30(10):1622–9.
38. Racle J, de Jonge K, Baumgaertner P, Speiser DE, Gfeller D. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *Elife*. 2017. <https://doi.org/10.7554/eLife.26476>.
39. Hong W, Liang L, Gu Y, Qi Z, Qiu H, Yang X, Zeng W, Ma L, Xie J. Immune-Related lncRNA to Construct Novel Signature and Predict the Immune Landscape of Human Hepatocellular Carcinoma. *Mol Ther Nucleic Acids*. 2020;22:937–47.
40. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS ONE*. 2014;9(9):e107468.
41. Hugo W, Zaretsky JM, Sun L, Song C, Moreno BH, Hu-Lieskovan S, Berent-Maoz B, Pang J, Chmielowski B, Cherry G, et al. Genomic and transcriptomic features of response to anti-PD-1 therapy in metastatic melanoma. *Cell*. 2016;165(1):35–44.
42. Lu X, Paliogiannis P, Calvisi DF, Chen X. Role of the mammalian target of rapamycin pathway in liver cancer: from molecular genetics to targeted therapies. *Hepatology*. 2021;73(Suppl 1):49–61.
43. Jang DK, Lee YG, Chan Chae Y, Lee JK, Paik WH, Lee SH, Kim YT, Ryu JK. GDC-0980 (apitolisib) treatment with gemcitabine and/or cisplatin synergistically reduces cholangiocarcinoma cell growth by suppressing the PI3K/Akt/mTOR pathway. *Biochem Biophys Res Commun*. 2020;529(4):1242–8.
44. Vagin VV, Sigova A, Li C, Seitz H, Gvozdev V, Zamore PD. A distinct small RNA pathway silences selfish genetic elements in the germline. *Science*. 2006;313(5785):320–4.
45. Han YN, Li Y, Xia SQ, Zhang YY, Zheng JH, Li W. PIWI proteins and PIWI-interacting RNA: emerging roles in cancer. *Cell Physiol Biochem*. 2017;44(1):1–20.
46. Mentis AA, Dardiotis E, Romas NA, Papavassiliou AG. PIWI family proteins as prognostic markers in cancer: a systematic review and meta-analysis. *Cell Mol Life Sci*. 2020;77(12):2289–314.
47. Iliiev R, Stanik M, Fedorko M, Poprach A, Vychytilova-Faltejskova P, Slaba K, Svoboda M, Fabian P, Pacik D, Dolezel J, et al. Decreased expression levels of PIWIL1, PIWIL2, and PIWIL4 are associated with worse survival in renal cell carcinoma patients. *Oncotargets Ther*. 2016;9:217–22.
48. Li W, Martinez-Useros J, Garcia-Carbonero N, Fernandez-Aceñero MJ, Orta A, Ortega-Medina L, Garcia-Botella S, Perez-Aguirre E, Diez-Valladares L, Celdran A, et al. The Clinical Significance of PIWIL3 and PIWIL4 Expression in Pancreatic Cancer. *J Clin Med*. 2020. <https://doi.org/10.3390/jcm9051252>.
49. Mishra NK, Niu M, Southeikal S, Bajpai P, Elkholy A, Manne U, Guda C. Identification of prognostic markers in cholangiocarcinoma using altered dna methylation and gene expression profiles. *Front Genet*. 2020;11:522125.
50. Henriques T, Scruggs BS, Inouye MO, Muse GW, Williams LH, Burkholder AB, Lavender CA, Fargo DC, Adelman K. Widespread transcriptional pausing and elongation control at enhancers. *Genes Dev*. 2018;32(1):26–41.
51. Shetty A, Kallgren SP, Demel C, Maier KC, Spatt D, Alver BH, Cramer P, Park PJ, Winston F. Spt5 plays vital roles in the control of sense and antisense transcription elongation. *Mol Cell*. 2017;66(1):77–88.e75.
52. Chen R, Zhu J, Dong Y, He C, Hu X. Suppressor of Ty homolog-5, a novel tumor-specific human telomerase reverse transcriptase promoter-binding protein and activator in colon cancer cells. *Oncotarget*. 2015;6(32):32841–55.

53. Lone BA, Ahmad F, Karna SKL, Pokharel YR. SUPT5H post-transcriptional silencing modulates pin1 expression, inhibits tumorigenicity, and induces apoptosis of human breast cancer cells. *Cell Physiol Biochem*. 2020;54(5):928–46.
54. Guo H, Cai J, Wang X, Wang B, Wang F, Li X, Qu X, Kong X, Gao Y, Wu H, et al. Prognostic values of a novel multi-mRNA signature for predicting relapse of cholangiocarcinoma. *Int J Biol Sci*. 2020;16(5):869–81.
55. Xie X, Wang Y, Zhang S, Li J, Yu Z, Ding X, Ye L, Gong P, Zhu Q, Li J, et al. A novel five-lncRNA signature panel improves high-risk survival prediction in patients with cholangiocarcinoma. *Aging (Albany NY)*. 2021;13(2):2959–81.
56. Mohibi S, Chen X, Zhang J. Cancer the RBP'etics-RNA-binding proteins as therapeutic targets for cancer. *Pharmacol Ther*. 2019;203:107390.
57. Lei X, Lei Y, Li JK, Du WX, Li RG, Yang J, Li J, Li F, Tan HB. Immune cells within the tumor microenvironment: biological functions and roles in cancer immunotherapy. *Cancer Lett*. 2020;470:126–33.
58. Zhang X, Shi M, Chen T, Zhang B. Characterization of the immune cell infiltration landscape in head and neck squamous cell carcinoma to aid immunotherapy. *Mol Ther Nucleic Acids*. 2020;22:298–309.
59. Zuo S, Wei M, Wang S, Dong J, Wei J. Pan-cancer analysis of immune cell infiltration identifies a prognostic immune-cell characteristic score (ICCS) in lung adenocarcinoma. *Front Immunol*. 2020;11:1218.
60. Wang S, Zhang Q, Yu C, Cao Y, Zuo Y, Yang L. Immune cell infiltration-based signature for prognosis and immunogenomic analysis in breast cancer. *Brief Bioinform*. 2021;22(2):2020–31.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

