Correlation of Hepatocellular Carcinoma Stemness Indices and Clinical Characteristics and Prognosis

Juan Li (fcclij2@zzu.edu.cn)
Zhengzhou University First Affiliated Hospital https://orcid.org/0000-0002-0817-6516
Chunting Zhang
Zhengzhou University First Affiliated Hospital
Xin Yuan
Zhengzhou University First Affiliated Hospital
Zujiang Yu
Zhengzhou University First Affiliated Hospital

Primary research

Keywords: hepatocellular carcinoma, cancer stem cell, mRNAsi, mDNAsi, prognosis

Posted Date: February 20th, 2020

DOI: https://doi.org/10.21203/rs.2.24061/v1

Abstract

Background

Recent studies have shown that cancer stem cell (CSC) is related to the occurrence and development of hepatocellular carcinoma(HCC), but the mechanism has not yet been elucidated. Therefore, it is necessary to explore the relationship between HCC stem cells and the survival time of HCC patients and its mechanism to guide the treatment.

Methods

We downloaded the RNA-Seq data and clinical information from The Cancer Genome Atlas (TCGA) and International Cancer Genome Consortium (ICGC). The mRNA gene expression-based stemness index (mRNAsi) and DNA methylation-based stemness index (mDNAsi) were calculated through one-class logistic regression (OCLR). By applying the univariable Cox regression analysis, we found that mRNAsi and mRNAsi were significantly correlated with overall survival (OS). Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO) analysis were used to seek for different genes, and searched for hub genes through protein-protein interaction (PPI) analysis. The spearman's rank correlation coefficient test was applied to analyze the relationship between these hub genes and the stemness indices.

Results

The mRNA gene expression-based stemness index (mRNAsi) and DNA methylation-based stemness index (mDNAsi) levels of HCC samples from TCGA and ICGC were significantly negatively related to clinical characteristics and OS. Analysis of differentially expressed genes and PPI revealed that SNAP25, KPT19, GABBR1 and EPCAM were significantly negative correlation with mDNAsi.

Conclusion

Our new model based on stemness indices-related genes was available for predicting prognosis. The SNAP25, KPT19, GABBR1 and EPCAM were potentially therapeutic targets for HCC patients.

Background

Hepatocellular carcinoma (HCC) is the sixth common cancer and the third most frequent cause of cancer death [1]. Despite surgical resection provides the most potential cure for the early stage HCC, the 5-year survival rate of HCC is less than 20%, mainly due to limited effective interventions for advanced HCC [2, 3]. However, own to the absence of specific clinical symptoms in early stages, most patients with HCC are diagnosed in an advanced stage with poor prognosis [3, 4]. But the specific pathogenesis of HCC is still not clear [5]. Thus, it is essential to further explore the development and progression mechanism of HCC to achieve early diagnosis and target treatment.

The emergence of the concept of cancer stem cells (CSCs) allows people to know that cancer cells are not exactly the same, but there are many types of cells [6, 7]. Therefore, malignant solid tumors consist of a highly heterogeneous cancer stem cell population. And CSCs illustrate that a small number of cancer cells can self-replicate as stem cells and maintain tumor characteristics, just as normal stem cells renew and maintain our organs and tissues [8, 9]. Growing evidence showed that CSCs are original cells of tumor and play crucial roles in tumorigenic, metastasis, and recurrence, and resistance to chemotherapy and radiation therapy [10]. However, the potential molecular mechanism and signaling pathway of CSCs in cancer, including HCC, are not clear.

Increasing studies have shown that CSCs are therapeutic targets of cancers [11]. Yang et al. found that HGF/c-Met promoted the enrichment of renal carcinoma cancer stem cells, which may be one of the possible mechanisms of metastasis and recurrence, and may be a new therapeutic target in renal cell carcinoma [12]. Increasing evidence has suggested that clarifying the specific source of CSCs can provide a novel method for individualized treatment [13]. However, little is known about the expression pattern and the specific mechanism of CSCs in HCC.

In this study, we built a model based on the sequencing data of stem cells in the Progenitor Cell Biology Consortium (PCBC) database to predict the stemness indices in tumor samples. Besides, we utilized TCGA cohort to make predictions, and validated our findings by an independent dataset. Furthermore, we analyzed the relationship between the stemness indices, clinical characteristics and OS of patients, and searched different genes between high and low stem cell factor levels samples. And PPI analysis was performed to find the hub genes, which might provide potential target for stem cell treatment of HCC.

Methods

Data Procession

In this study, the following databases were used: The Cancer Genome Atlas (TCGA, https://cancergenome.nih.gov/) database, International Cancer Genome Consortium (ICGC, https://icgc.org/) database, Progenitor Cell Biology Consortium (PCBC, https://progenitorcells.org/frontpage) database and The STRING (https://string-db.org/) database. The expression data (syn2701943) of the pluripotent stem cell samples including embryonic stem cells (ESC) and induced pluripotent stem cells (iPSC) were downloaded from the PCBC using the R package synapse (v0.6.61). The RNA-Seq data and clinical follow-up information of 424 samples were downloaded in TCGA at the same time. The same information contains 203 samples were downloaded by ICGC, which used as validation set. The clinical follow-up information included gender, age, pathological stage, TNM staging, whether HBV infection, HCV infection, immune cell score and iCluster typing. The data was processed as described previously [14, 15].

Calculation Of Stem Cell Index Based On OCLR Method

We used the expression data of the pluripotent stem cell samples (ESC and iPSC) from the PCBC database, and used the one-class logistic regression (OCLR) method to predict and calculate the stem cell

index. The mothod of calculation was the same as before [16, 17]. Finally, there are 78 SC samples in total, 12998 expression profiles of mRNA genes per sample. In the calculation of mDNAsi, 99 samples with 219 stem cell probes signature in the PCBC database were divided into two groups, of which 44 samples were stem cell samples and 55 samples were non-stem cell samples. For the expression profile and methylation data obtained, centralized by the average value for each sample, and finally calculated the mRNAsi and mDNAsi weight vector of each gene by OCLR method in the R package gel net (v1.2.1).

For the expression profile and methylation data of TCGA and ICGC tumor samples, we calculated the Spearman correlation coefficient of the gene expression profile and methylation data of each sample gene with the model's gene weight vector. Perform a linear conversion on the Spearman correlation coefficient obtained from the sample. The conversion methods were utilized to identify the minimum Spearman correlation coefficient from the Spearman correlation coefficient of each sample, and then divide it by the maximum value. The value obtained is the stem cell index-mRNAsi and mDNAsi of each sample, the stem cell index distribution range [0, 1].

Analyzing And Screening Of Differentially Expressed Genes

We used the R package DESeq2 (v1.24.0) to perform a differential analysis on the RNA-seq data of the mRNAsi-high/ mDNAsi-high and mRNAsi-low/ mRNAsi-low samples, and calculated the differentially expressed genes (DEGs) between the mRNAsi-high/ mDNAsi-high sample and the mRNAsi-low/ mRNAsi-low samples. Differential genes were filtered according to the threshold false discovery rate (FDR) < 0.05 and |log2FC| > 1.

Enrichment Analysis Of Gene Function And Pathways

Kyoto Encyclopedia of Genes and Genomes (https://www.kegg.jp/ or https: //www.genome. jp/ keg /, KEGG) pathway enrichment analysis, which revealed a network of metabolism, signaling, and other molecular interactions [18]. Gene Ontology (GO) function enrichment analysis is an international standard classification system for gene functions, including Biological Process (BP), Cellular Components (CC), and Molecular Function (MF) [19]. Through these three functional categories, gene functions were defined and described in many aspects. To examine the function of stemness indices differential genes, we performed KEGG pathway analysis and GO function enrichment analysis on the differential genes of mRNAsi and mDNAsi through the R package cluster Profiler (v3.14.0). The significance level was defined as FDR < 0.05.

PPI Network Analysis

The STRING database was used to analyze the PPI network of differential genes, and the degree algorithm of the cytoHubba plug-in in Cityscape software was applied to identify the hub genes. PPI network of organism serves as the backbone of its signaling pathways that mediate cellular responses to the environment and heredity [20]. PPI is commonly used to elucidate the molecular basis of disease and provide detailed knowledge about proteins and their interactions, which can be used to suggest and

improve the diagnosis, prevention, and treatment of a disease [21]. PPI network analysis provides information about genes and proteins shared in disease and describes interactions [22].

Statistics

A t test was used for continuous variables. The statistical analyses were performed using IBM SPSS 24.0. All statistical tests were two-sided, and the significance level was defined as $P < 0.05$. Spearman's rank correlation coefficients were tested to analyze the correlation between two variables. Kaplan-Meier survival curves were generated after stratification of the data and compared with a log-rank test.

Results

mRNA expression- and DNA methylation-based stemness indices in HCC

Firstly, to explore the relationship between the index distributions of stem cells and clinical characteristics. We ranked the HCC samples according to their mRNA gene expression-based stemness index (mRNAsi) and DNA methylation-based stemness index (mDNAsi) values and tested whether any clinical feature was correlated with low or high stemness index. The mRNAsi and mDNAsi exponential distribution of TCGA HCC samples were shown in Fig. 1a and Fig. 1b, respectively. The mRNAsi exponential distribution of ICGC HCC samples was shown in Fig. 1c.

Relationship Between Stem Cell Index And Clinical Characteristics

To further explore the relationship between mRNAsi and mDNAsi indexes and clinical features of tumor samples, such as: gender, age, pathological stage, whether HBV or HCV infection, immune cell score,TNM stage and iCluster typing, we performed a linear regression test on the above characteristics. As shown in Table 1, the mRNAsi of TCGA samples were significantly related to pathological stage, HBV infection, immune cell score, and iCluster typing ($P < 0.05$). And the mDNAsi was remarkably associated with immune cell score and iCluster typing ($P < 0.05$). Furthermore, we analyzed the data from ICGC cohort, which indicated that the mRNAsi was significantly associated with TNM stage (Table 2, $P < 0.05$).

Table 1
Correlation between mRNAsi and mDNAsi indices of TCGA liver cancer samples and their characteristic information and molecular typing

	Percent/Average	p(mRNAsi)	p(mDNAsi)
Sex	Male:65.6%	0.80	0.12
	Female:34.4%		
Age	60.13	0.47	0.23
Grade	G1:17.8%	3.71E-05	0.67
	G2:48.3%		
	G3:33.9%		
HBV	23.00%	3.04E-04	0.47
HCV	18.60%	0.76	0.15
Immune score	166.12	1.2E-03	0.018
iCluster	iCluster1:35.5%	1.29E-07	5.64E-31
	iCluster2:30.1%		
	iCluster3:34.4%		

Table 2
Correlation between the mRNAsi index of ICGC liver cancer samples and the characteristic information and molecular typing of the samples

	Percent/Average	P (mRNAsi)
Sex	Male:75.4%	0.12
	Female:24.6%	
Age	67	0.2
TNMStage	I:16.3%	9.09E-03
	II:47.3%	
	III:29.1%	
	IV:7.4%	
HBV	23.00%	0.81
HCV	18.60%	0.64
Immunescore	166.12	0.97

Table 3
HR analysis and optimal threshold for TCGA stem cell index grouping

iCluster	mRNASi					
	hazard.ratio				Cut P	
	Hazard ratio	lower	upper	P	Cut	P (Cut)
C1	1.80	0.96	3.40	6.77E-02	0.37	0.23
C2	3.61	1.20	10.84	2.21E-02	0.35	0.08
C3	1.60	0.79	3.23	0.19	0.43	0.77
All	1.68	1.11	2.53	0.0145	0.38	0.08
iCluster	mDNAsi					
	Hazard ratio				Cut P	
	Hazard ratio	lower	upper	P	Cut	P (Cut)
C1	0.69	0.37	1.31	0.26	0.12	0.87
C2	0.52	0.20	1.34	1.75E-01	0.20	0.67
C3	1.85	0.91	3.76	0.09	0.30	0.48
All	0.62	0.39	0.98	4.04E-02	0.12	0.51

In TCGA, the mRNAsi level was significantly higher in HBV-infected samples compared with uninfected samples (Fig. 2a, $P < 0.05$), and the mRNAsi score of iCluster 3 samples was significantly higher than that of iCluster 1 and iCluster 2 (Fig. 2b, c, $P < 0.05$). In addition, the patients with late pathological stage have the higher mRNAsi expression. (Fig. 2d, e, $P < 0.05$). The iCluster typing was significantly correlated with mDNAsi. The mDNAsi score was lowest in iCluster1 and highest in iCluster3.While the HBV-infected and pathological stage had no obviously correlation with mDNAsi. (Supplementary Fig. 1a-e, $P < 0.05$). In ICGC, TNM was significantly related to the mRNAsi, while HBV-infected subgroup exhibited no significantly relation with the mRNAsi. The samples with higher mRNAsi index probably in higher TNM stage (Fig. 2f, g, $P < 0.05$). The immune score was negatively related to mRNAsi and mDNAsi in TCGA (Fig. 2h, Supplementary Fig. 1f, $P = 0.0012$). Furthermore, the correlation between the mRNAsi and mDNAsi indexes of tumor samples were validated and indicating that these two subtypes of indexes were significantly positively correlated ($cor = 0.30$, Fig. 2i, $P = 0.00012$). Taken together, the stemness indices were negatively related with clinical characteristics, samples with worse clinical characteristics tended to have higher stemness indices, demonstrating that the stem cell indexes could play essential roles in tumor progression.

Correlations Of Stemness Indices With OS In HCC

To better understand the relationship between stemness indices and prognosis, we compared differences of OS with clinical features such as gender, age, pathological stage, iCluster typing, immune cell scores, and HBV or HCV infection among various samples. The results illustrated that clinical factors such as gender, age had no significant correlation with OS (Fig. 3a, b, P = 0.26, P = 0.6). And the patients over survival presented no obviously relation with HCV infection, pathological stage (Fig. 3c, d, P = 0.42, P = 0.25) and iCluster typing (Fig. 3e, P = 0.22).There were significant differences in OS of samples grouped based on HBV infection and immune cell scoring. The high-immune score group demonstrated higher overall survival rate compared with low-immune score group (Fig. 3f, P = 0.012). And the survival probability of group of HBV infected was lower than the group of not infected (Fig. 3g, P = 0.036). The OS of mRNAsi-low group was higher than the mRNAsi-high group, and the mDNAsi-low group was higher than the mDNAsi-high group (Fig. 3h, Supplementary Fig. 2a, P = 0.014). Then the samples grouped based on the stemness indices mRNAsi and mDNAsi were identified according to iCluster typing. Only in the iCluster2 sample group, there was a significant difference of OS in samples grouped by the mRNAsi (Fig. 3i, j, k, Supplementary Fig. 2b, c, d, P = 0.064, P = 0.015, P = 0.18). Shorter and fewer deaths made it difficult to distinguish.

Then we performed further verification in the ICGC database indicating that the OS had no correlation with gender, age (Fig. 4a, b, P = 0.077, P = 0.2), and also had no relevant with immune cell scores, HBV infection (Fig. 4c, d, P = 0.15, P = 0.88). The overall survival rate of HCV infection subgroup was lower and the higher TNM stage had lower OS (Fig. 4e, f, P = 0.033, P < 0.05). And the OS of mRNAsi-low group was higher than mRNAsi-high group (Fig. 4g, P < 0.001). All these hinted, the stemness indices negatively related to OS and affected the prognosis of patients, which may provide a new objective criterion for evaluating the prognosis of HCC patients.

Screening The Stem Cell-related Genes In HCC

Based on the above results, we further analyzed the differentially expressed genes (DEGs). In TCGA cohort, a total of 136 differential genes were filtered according to the threshold, among which the mRNAsi-high group was down-regulated by 110 genes and up-regulated by 26 genes compared to the mRNAsi-low group. The DEGs were mainly down-regulated differentially expression (Fig. 5a). A total of 569 differential genes were selected after filtering, of which mDNAsi-high group were down-regulated by 546 genes and up-regulated by 23 genes compared to mDNAsi-low group. The DEGs were mainly down-regulated differentially expression (Fig. 5b). Among the differential genes in the two sets, mRNAsi and mDNAsi were both down-regulated genes with intersect into four genes, respectively APOBEC3C, C1orf116, GABBR2 and TFCP2L1. In ICGC cohort, the volcano map of DEGs showed there were 190 differential genes. Among them, mRNAsi-high group have 16 genes down-regulated and 174 genes up-regulated compared to mRNAsi-low group, which were major up-regulated differential expression (Fig. 5c). The intersection of the up-regulated genes with the TCGA samples based on the mDNAsi grouping was ALDH1A3. These up- or down-regulated differential genes reflect different gene expressions in HCC samples with high and low stem cell indices, which provides new ideas for the further exploration of HCC pathological mechanisms.

Functional Analysis Of Differentially Expressed Genes

To further explore the function of differentially expressed genes and their roles in tumor genesis, metastasis, recurrence, and drug resistance, a functional analysis of them were performed. GO function annotations of 136 differential genes of TCGA tumor samples grouped by mRNAsi showed without significant enrichment (FDR < 0.05). GO function annotations of 569 differential genes of TCGA tumor samples identified by mDNAsi indicating that the pathway “metal ion trans-membrane transporter activity” and “monovalent inorganic action trans-membrane transporter activity” were significantly enriched (Fig. 5d, FDR < 0.05). KEGG pathway enrichment analysis of mRNAsi differential genes found no significant enriched pathway (FDR < 0.05). For the KEGG pathway enrichment analysis, our results demonstrated the pathway “Camp signaling pathway” was significantly enriched (Fig. 5e, FDR < 0.05). For the GO function annotations of 569 differential genes obtained by ICGC tumor samples based on mRNAsi grouping, the pathway “extracellular matrix structural constituent” and “glycosaminoglycan binding” were significantly enriched (Fig. 5f, FDR < 0.05). The KEGG pathway enrichment analysis of mRNAsi differential genes found three significantly enriched pathways. “PI3K-Akt signaling pathway” was significantly enriched (Fig. 5g, FDR < 0.05). Taken together, we found the DEGs of high and low stemness indices always enriched on pathway correlated with tumor formation, metastasis and recurrence, which help to understand the tumor genesis mechanism and guide the targeted therapy of tumors.

PPI analysis and hub genes analysis of differentially expressed genes

To further find out hub genes related to stemness indices, PPI analysis of differentially expressed genes were performed. 10 genes (SNAP25, EPCAM, CALB2, SOX2, KRT19, GABBR1, MUC1, AFP, GRIN1, and GAD1) were selected as hub genes (Fig. 6a). The correlation between these hub genes and the mDNAsi index were performed. The results suggested that SNAP25, KPT19, GABBR1 and EPCAM showed a significant negative correlation with mDNAsi (Fig. 6b-e, P < 0.0001, P = 0.0023, P = 0.039, P = 0.0006), while the other six hub genes had no prominent connection (Fig. 6f-k, P = 0.062, P < 0.039, P = 0.7, P = 0.21, P = 0.12, P = 0.94). In general, 10 genes were found as hub genes related to stemness indices through PPI analysis. And 4 of 10 hub genes were significantly negatively correlated with mDNAsi. It suggested that the tumor stemness indices could be used as a new objective indicator to predict the prognosis of patients with HCC. SNAP25, KPT19, GABBR1 and EPCAM might key genes for HCC occurrence, metastasis and recurrence, which provided new ideas for targeted treatment of HCC.

Discussion

In this study, we found that the stemness indices were related to pathological stage, HBV infection, immune cell score, icluster typing and TNM stage. Several studies have demonstrated that the stem cell index is associated with pathological characters in many cancers, which was consistent with our findings in HCC. For example, Pan et al. clarified that corrected mRNAsi in bladder cancer increased as tumor stage increased, with T3 having the highest stem cell characteristics. Lower corrected mRNAsi scores had

better overall survival and treatment outcome [23]. A recent study reported the cancer stem cell marker OCT4 expression was correlated with poor differentiation, tumor size, and N stage in patients with rectal cancer, and it could be an independent prognostic biomarker [24]. All of these results suggested that the stemness indices could play essential roles in tumor progression.

Increasing number of studies had revealed that the stem cell index was significantly related to clinical prognosis of patients. In our study, we constructed a novel model to predict the over survival in patients with HCC, and detected the relationship between HCC stemness indices and clinical characteristics and OS. Our results showed that patients with worse clinical phenotype usually possessed higher stemness indices and shorter OS time. Consistent with our findings, previous study indicated that overexpression of the cancer stem cell marker CD133 confers a poor prognosis in invasive breast cancer [25]. Qin et al. found that mRNAsi is an independent prognostic factor in lung squamous cell carcinoma [26]. Zhao et al. found that TBX21 predicts prognosis of patients and drives cancer stem cell maintenance via the TBX21–IL-4 pathway in lung adenocarcinoma. And TBX21 could serve as a novel predictive biomarker and therapeutic target for patients [27]. Additionally, Lian et al. expounded that the prognostic signature based on mRNAsi may contribute to personalized prediction of sonic hedgehog medulloblastoma prognosis and act as a potential biomarker for prognostication and response to differentiation therapies in clinical practice [17]. These findings demonstrated that stemness indices was a reliable predictor of the prognosis of HCC.

Furthermore, we investigated the differentially expressed genes of stemness indices, and analyzed the function of differential genes. The pathways “metal ion trans-membrane transporter activity” “monovalent inorganic action trans-membrane transporter activity” “cAMP signaling pathway” “extracellular matrix structural constituent” “glycosaminoglycan binding” and “PI3K/Akt signaling pathway” were significantly enriched. Roberto et al. found that synergistic inhibition of HCC and liver cancer stem cell proliferation by targeting RAS/RAF/MAPK and WNT/β-Catenin pathways [28]. It was reported that MicroRNA-28-5p regulates liver cancer stem cell expansion via IGF-1 pathway [29]. Si et al. indicated that the miR219/E-cadherin axis as a potential therapeutic target against liver CSCs and a potential predictor for sorafenib treatment of HCC patients [30]. Based on these results, the differentially expressed genes related to stem cell were correlated with differentially tumor genesis-related pathways, which revealed potential mechanism of CSCs related genes in HCC.

Additionally, PPI were performed to seek for hub genes, and found that SNAP25, KPT19, GABBR1 and EPCAM are significantly correlation with mDNAsi. Similar to our results, previous study showed that SENP1 activity sustains cancer stem cell in hypoxic HCC [31]. Li et al. clarified activated STAT3 plays a pivotal role in holding cancer stemness of HCC CSCs [32]. Ritu et al. elucidated that miR-26b-5p imparts metastatic properties and helps in maintenance of Ep + CSCs via HSPA8. Thus, miR-26b-5p and HSPA8 could serve as molecular targets for selectively eliminating the Ep + CSC population in human HCC [33]. Therefore, all these findings further confirmed that hub genes related to HCC stem cells could help understand the pathogenesis, indicating a novel idea for targeted therapy in HCC.

Conclusions

This might advance the development of objective diagnostic tools for quantitating HCC stemness and we can make prognostic judgments and predict treatment responses based on stem cell index information, and provide new ideas for our clinical treatment, develop personalized treatment methods. Further exploration of stemness indices-related genes on tumor proliferation and metastasis is still in need.

Abbreviations

CSC: cancer stem cell; HCC: hepatocellular carcinoma; mRNAsi: mRNA gene expression-based stemness index; mDNAsi: DNA methylation-based stemness index; TCGA: The Cancer Genome Atlas; ICGC: International Cancer Genome Consortium; OCLR: one-class logistic regression; KEGG: Kyoto Encyclopedia of Genes and Genomes; GO: Gene Ontology; PPI: protein-protein interaction; OS: overall survival; PCBC: Progenitor Cell Biology Consortium; ESC: embryonic stem cells; iPSC: induced pluripotent stem cells; FDR: false discovery rate; BP: Biological Process; CC: Cellular Components; MF: Molecular Function; DEGs: differentially expressed genes.

Declarations

Acknowledgements

We earnestly appreciate all the participators in our work. We thank the patients and investigators who participated in TCGA, PCBC and ICGC for providing data.

Authors' contributions

JL and ZJY: study concepts, study design, literature research, manuscript preparation and editing; CTZ: definition of intellectual content, literature research, manuscript preparation and editing; XY: data acquisition, literature research and manuscript writing. All authors read and approved the final manuscript.

Funding

This study was supported by funds from the National Natural Science Foundation of China (81702927).

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Competing interests

The authors have no conflicts of interest to report.

References

1. Wang D, Hu X, Xiao L, Long G, Yao L, Wang Z, Zhou L: **Prognostic Nutritional Index and Systemic Immune-Inflammation Index Predict the Prognosis of Patients with HCC**. *Journal of Gastrointestinal Surgery* 2020.
2. Li X, Li C, Zhang L, Wu M, Cao K, Jiang F, Chen D, Li N, Li W: **The significance of exosomes in the development and treatment of hepatocellular carcinoma**. *Molecular Cancer* 2020, **19**(1).
3. Jiao J, González Á, Stevenson HL, Gagea M, Sugimoto H, Kalluri R, Beretta L: **Depletion of S100A4+ stromal cells does not prevent HCC development but reduces the stem cell-like phenotype of the tumors**. *Experimental & Molecular Medicine* 2018, **50**(1):e422-e422.
4. Zelvyte I, Stevens T, Westin U, Janciauskiene S: **alpha1-antitrypsin and its C-terminal fragment attenuate effects of degranulated neutrophil-conditioned medium on lung cancer HCC cells, in vitro**. *Cancer Cell Int* 2004, **4**(1):7.
5. Ding B, Lou W, Liu J, Li R, Chen J, Fan W: **In silico analysis excavates potential biomarkers by constructing miRNA-mRNA networks between non-cirrhotic HCC and cirrhotic HCC**. *Cancer Cell Int* 2019, **19**:186.
6. Liu D, Hong Y, Li Y, Hu C, Yip TC, Yu WK, Zhu Y, Fong CC, Wang W, Au SK *et al*: **Targeted destruction of cancer stem cells using multifunctional magnetic nanoparticles that enable combined hyperthermia and chemotherapy**. *Theranostics* 2020, **10**(3):1181-1196.
7. Fundamental MLJ, pharmacology c: **Cancer stem cells as therapeutic targets of pancreatic cancer**. 2020.
8. Huang L, Lian J, Chen X, Qin G, Zheng Y, Zhang Y: **WASH overexpression enhances cancer stem cell properties and correlates with poor prognosis of esophageal carcinoma**. *Cancer Sci* 2017, **108**(12):2358-2365.
9. S H, S P, J A, S M, M M, F H, S G, A K, A A, APJCP VJJAPjocp: **Characterization and Validation of Hepatocellular Carcinoma (HCC) Xenograft tumor as a Suitable Liver Cancer Model for Preclinical Mesenchymal Stem Cell Studies**. 2018, **19**(6):1627-1631.
10. Abad E, Graifer D, Lyakhovich A: **DNA damage response and resistance of cancer stem cells**. *Cancer Lett* 2020, **474**:106-117.
11. Y G, B L, X Y, X S, J M, S L, Chemosphere ZDJ: **Bisphenol A and polychlorinated biphenyls enhance the cancer stem cell properties of human ovarian cancer cells by activating the WNT signaling**

pathway. 2019, **246**:125775.

12. Yang J, Yang L, Li S, Hu N: **HGF/c-Met Promote Renal Carcinoma Cancer Stem Cells Enrichment Through Upregulation of Cir-CCDC66**. *Technol Cancer Res Treat* 2020, **19**:1533033819901114.
13. Yen CH, Lai CC, Shia TH, Chen M, Yu HC, Liu YP, Chang FR: **Gynura divaricata attenuates tumor growth and tumor relapse after cisplatin therapy in HCC xenograft model through suppression of cancer stem cell growth and Wnt/beta-catenin signalling**. *J Ethnopharmacol* 2018, **213**:366-375.
14. Y H, Q D, J L, Q Z, X Y, M X, Aging GWJ: **Prediction of hepatocellular carcinoma prognosis based on expression of an immune-related gene set**. 2020, **12**(1):965-977.
15. Y H, C X, Y Y, J C, X C, F R, Z R, G C, management SRJC, research: **CD44 is overexpressed and correlated with tumor progression in gallbladder cancer**. 2018, **10**:3857-3865.
16. F J, A E, G M, S D, journal VCJTEr: **Transforming growth factor-beta1 is a potent inhibitor of secretory leukoprotease inhibitor expression in a bronchial epithelial cell line. Munich Lung Transplant Group**. 2000, **15**(6):1052-1057.
17. H L, YP H, YC Z, Y Z, S Y, QF L, BC W, JJ W, W M, J Y *et al*: **Integrative analysis of gene expression and DNA methylation through one-class logistic regression machine learning identifies stemness features in medulloblastoma**. 2019, **13**(10):2227-2245.
18. M K, Society SYJPsapotP: **KEGG Mapper for inferring cellular functions from protein sequences**. 2020, **29**(1):28-35.
19. X Z, X Z, Y L, Z C, Z Y, A X, international HJJBr: **Identification of Key Modules, Hub Genes, and Noncoding RNAs in Chronic Rhinosinusitis with Nasal Polyps by Weighted Gene Coexpression Network Analysis**. 2020, **2020**:6140728.
20. F Y, X P, L C, YH Z, T H, international CYJBr: **Analysis of Protein-Protein Functional Associations by Using Gene Ontology and KEGG Pathway**. 2019, **2019**:4963289.
21. D S, communications SRJN: **A systematic approach to orient the human protein-protein interaction network**. 2019, **10**(1):3015.
22. B R, N A-A, sciences M-HZJIjom: **Protein-Protein Interaction Network Analysis Reveals Several Diseases Highly Associated with Polycystic Ovarian Syndrome**. 2019, **20**(12).
23. S P, Y Z, X C, B W, oncology LBJFi: **Identification of Biomarkers for Controlling Cancer Stem Cell Characteristics in Bladder Cancer by Network Analysis of Transcriptome Data Stemness Indices**. 2019, **9**:613.
24. L Y, X G, journal HYJYm: **Correlation of Cancer Stem-Cell Markers OCT4, SOX2, and NANOG with Clinicopathological Features and Prognosis in Operative Patients with Rectal Cancer**. 2018, **59**(1):35-42.
25. C J, M A, S K, M A, IM M, S A-I, MS T, FQ G, SJ J, SG M *et al*: **Overexpression of the cancer stem cell marker CD133 confers a poor prognosis in invasive breast cancer**. 2019, **174**(2):387-399.
26. S Q, X L, Q Z, investigation ZWJC: **Co-Expression Network Analysis Identified Genes Associated with Cancer Stem Cell Characteristics in Lung Squamous Cell Carcinoma**. 2020, **38**(1):13-22.

27. S Z, W S, J Y, research WLJSc, therapy: **TBX21 predicts prognosis of patients and drives cancer stem cell maintenance via the TBX21-IL-4 pathway in lung adenocarcinoma**. 2018, **9**(1):89.
28. R G, E M, M S, MF D, C C, BT S, research GRJA: **Synergistic inhibition of HCC and liver cancer stem cell proliferation by targeting RAS/RAF/MAPK and WNT/β-catenin pathways**. 2014, **34**(4):1709-1713.
29. Q X, T H, P Y, R W, H L, J Z, international ZXJSc: **MicroRNA-28-5p Regulates Liver Cancer Stem Cell Expansion via IGF-1 Pathway**. 2019, **2019**:8734362.
30. A S, L W, K M, R Z, H J, Z L, Z C, X F, cycle HBJC: **miR-219 regulates liver cancer stem cell expansion via E-cadherin pathway**. 2019, **18**(24):3550-3561.
31. A C, M T, Gut PMJ: **SENP1 activity sustains cancer stem cell in hypoxic HCC**. 2017, **66**(12):2051-2052.
32. W L, Q Z, K C, Z S, J L, Q Y, complementary LJJB, medicine a: **2-Ethoxystypandrone, a novel small-molecule STAT3 signaling inhibitor from Polygonum cuspidatum, inhibits cell growth and induces apoptosis of HCC cells and HCC Cancer stem cells**. 2019, **19**(1):38.
33. R K, H H, A R, G R, SK S, Liver TNJLiojotIAftSot: **miR-26b-5p helps in EpCAM+cancer stem cells maintenance via HSC71/HSPA8 and augments malignant features in HCC**. 2019, **39**(9):1692-1703.

Supplemental Figure Legends

Supplementary Fig. 1 Relationship between stemness indices and clinical characteristics in TCGA. **a** mDNAsi has no significantly relation with HBV infection. **b** mDNAsi is highest in iCluster3 and lowest iniCluster1 in samples without HBV infection. **c** mDNAsi has no significantly relation with pathological grade in samples without HBV infection. **d** mDNAsi is highest in iCluster3 and lowest iniCluster1 in all TCGA samples. **e** mDNAsi has no significantly relation with pathological grade in all TCGA samples. **f** The immune score is inversely related to mDNAsi. $*P < 0.05$

Supplementary Fig. 2 Kaplan-Meier survival analyses of TCGA tumor samples. **a** The survival probability of the mDNAsi-low group is higher than the mRNAsi-high group in all TGCA samples. **b** The survival probability of the mDNAsi-low group and the mDNAsi-high group in iCluster1 has no significantly difference. **j** The survival probability of the mDNAsi-low group and the mDNAsi-high group in iCluster2 has no significantly difference. **k** The survival probability of the mDNAsi-low group and the mDNAsi-high group in iCluster3 has no significantly difference.

Figures

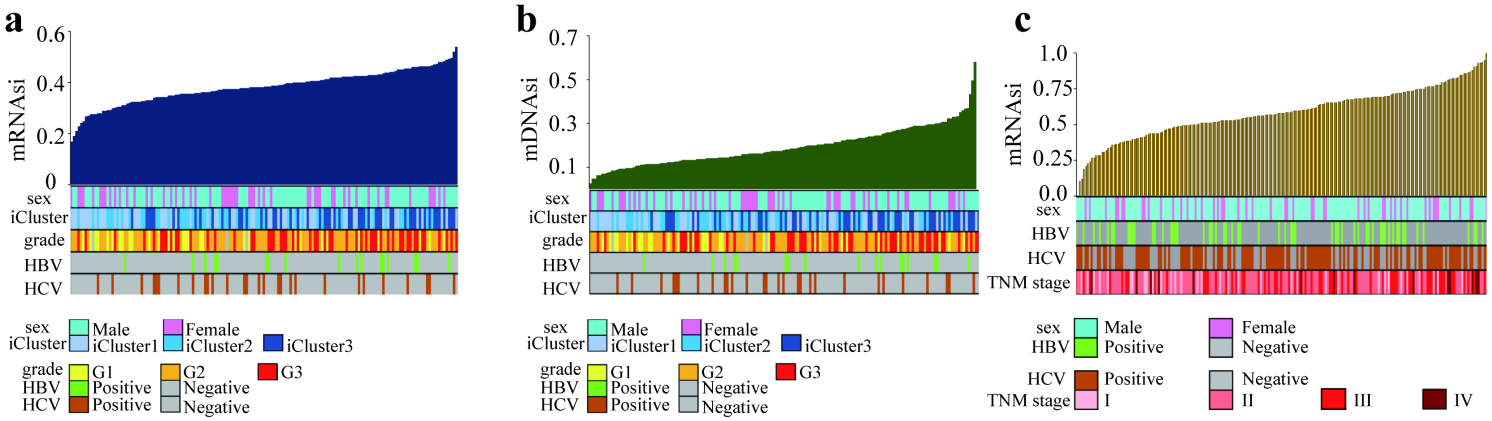

Figure 1

The mRNAsi and mDNAsi exponential distribution and clinical characteristic in HCC. a The mRNAsi exponential distribution of TCGA HCC samples. b The mDNAsi exponential distribution of TCGA HCC samples. c The mRNAsi exponential distribution of ICGC HCC samples.

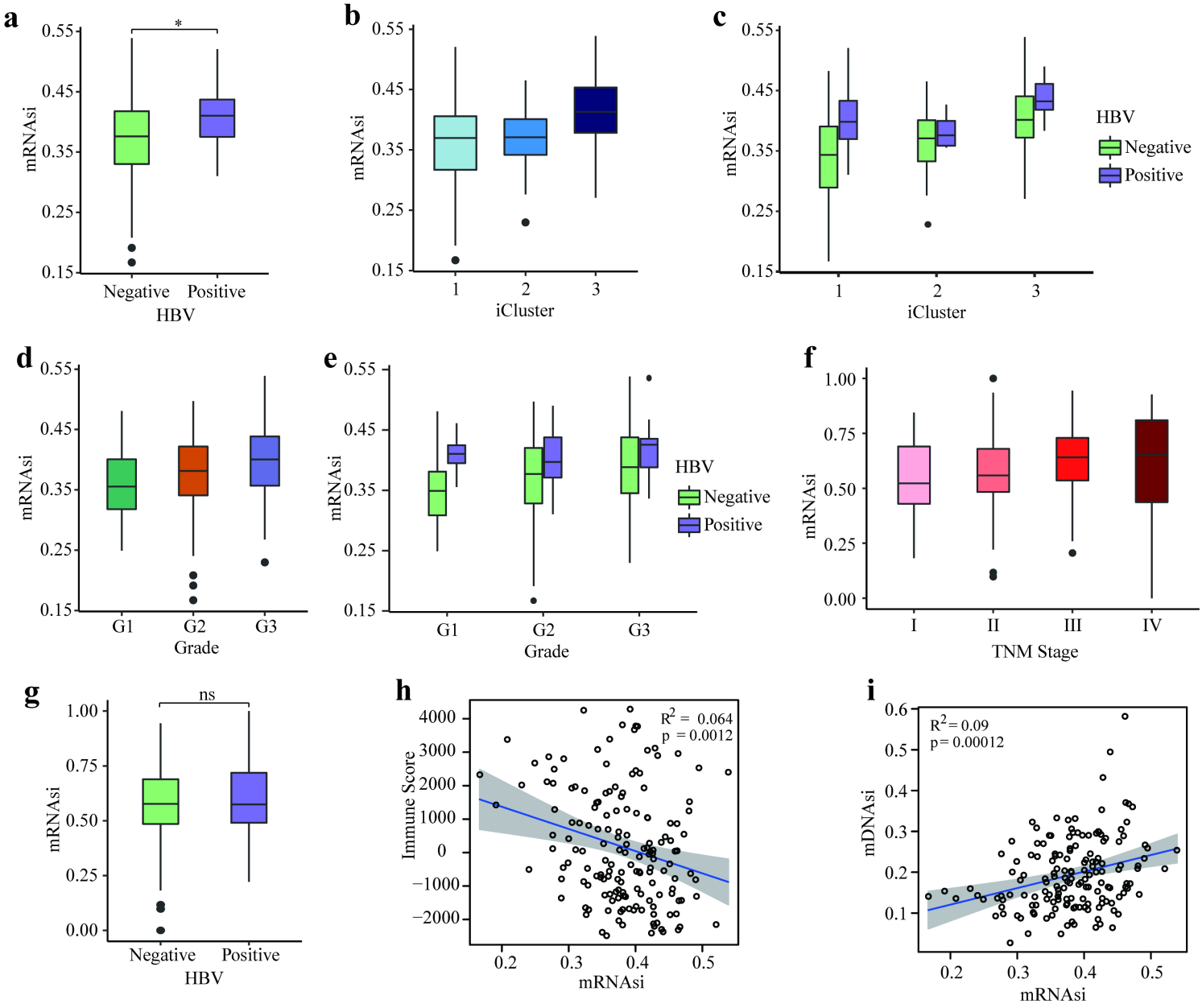

Figure 2

Relationship between mRNAsi and clinical characteristics. a mRNAsi is significantly higher in samples infected with HBV in TCGA. b mRNAsi is higher in iCluster 3 compare with iCluster 1 and iCluster 2 in all TCGA samples. c mRNAsi is the highest in iCluster 3 and the lowest in iCluster 1 in TCGA samples without HBV infected. d The mRNAsi is higher in the higher pathological grade in all TCGA samples. e The mRNAsi is higher in the higher pathological grade in TCGA samples without HBV infected. f The mRNAsi is higher in the higher TNM stage in ICGC samples. g mRNAsi and HBV infected have no significant collection. h The immune score is inversely related to mRNAsi. i mRNAsi and mDNAsi indices are positive correlation. *P < 0.05

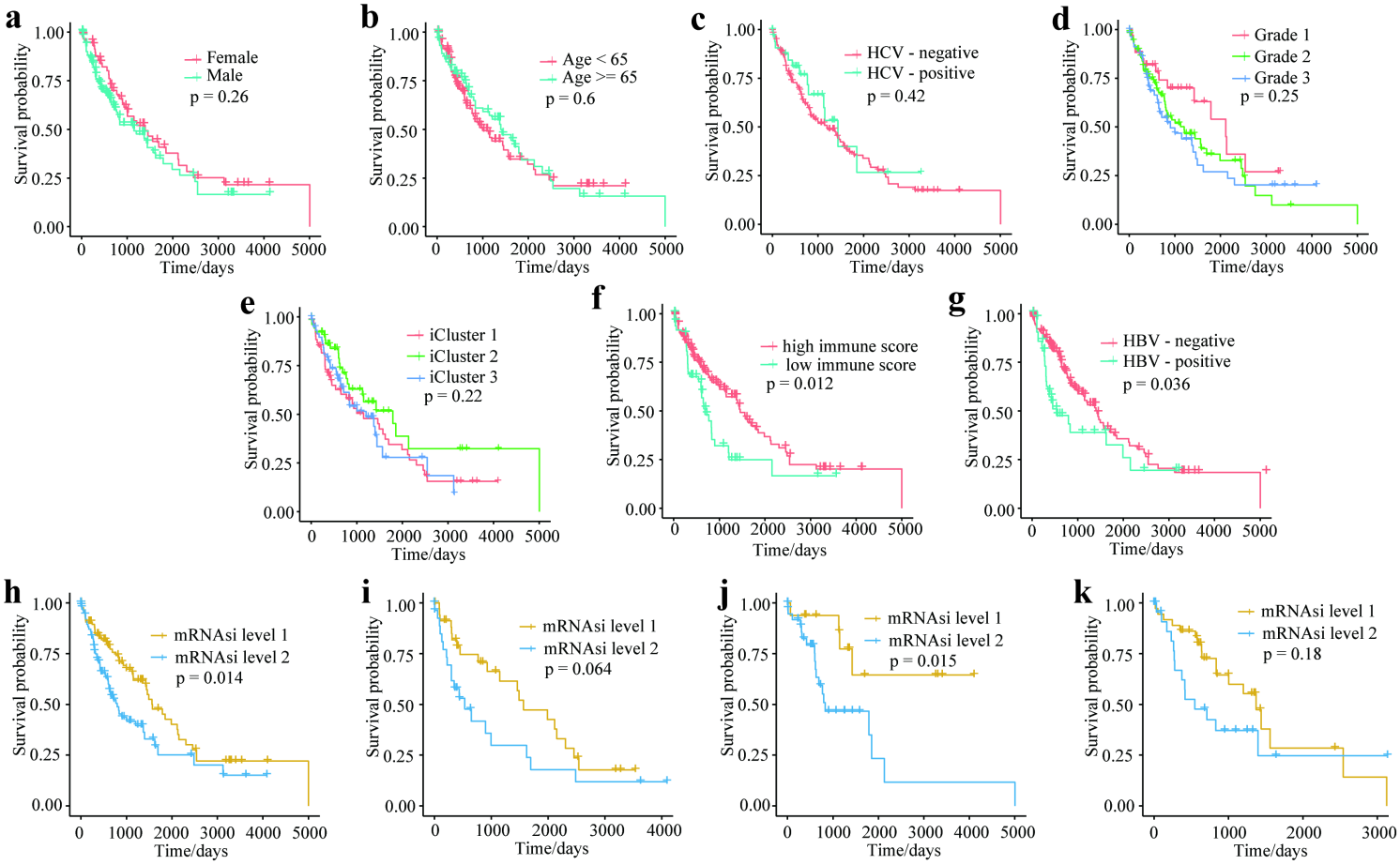

Figure 3

Kaplan-Meier survival analyses of TCGA tumor samples. a The survival probability of samples grouped by sex has no significantly difference. b The survival probability of samples grouped by age has no significantly difference. c The survival probability of samples grouped by HCV infection has no significantly difference. d The survival probability of samples grouped by iCluster has no significantly difference. e The survival probability of samples grouped by pathological stage has no significantly difference. f The high immune score group is higher than the low immune score group. g. The survival probability of group of HBV infected is lower. h The survival probability of the mRNAsi-low group is higher than the mRNAsi-high group in all TGCA samples. i The survival probability of the mRNAsi-low group and the mRNAsi-high group in iCluster1 has no significantly difference. j The survival probability of the mRNAsi-low group is higher than the mRNAsi-high group in iCluster2. k The survival probability of the mRNAsi-low group and the mRNAsi-high group in iCluster3 has no significantly difference.

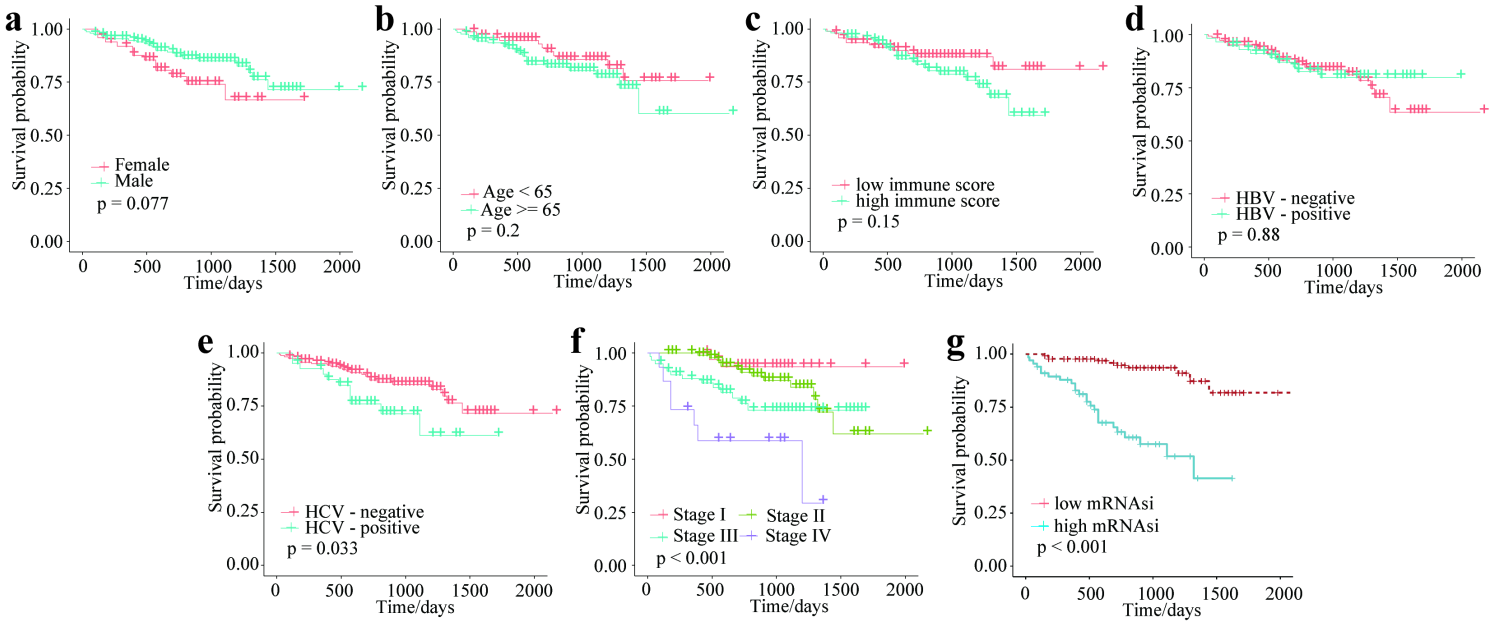

Figure 4

Kaplan-Meier survival analyses of ICGC tumor samples. a The survival probability of samples grouped by gender has no significantly difference. b The survival probability of samples grouped by age has no significantly difference. c The survival probability of samples grouped by immune cell scores has no significantly difference. d The survival probability of samples grouped by HBV infection has no significantly difference. e The survival probability of samples with HCV infected is lower. f The survival probability is lower in higher TNM stage. g The survival probability of the mRNAsi-low group is higher than the mRNAsi-high group.

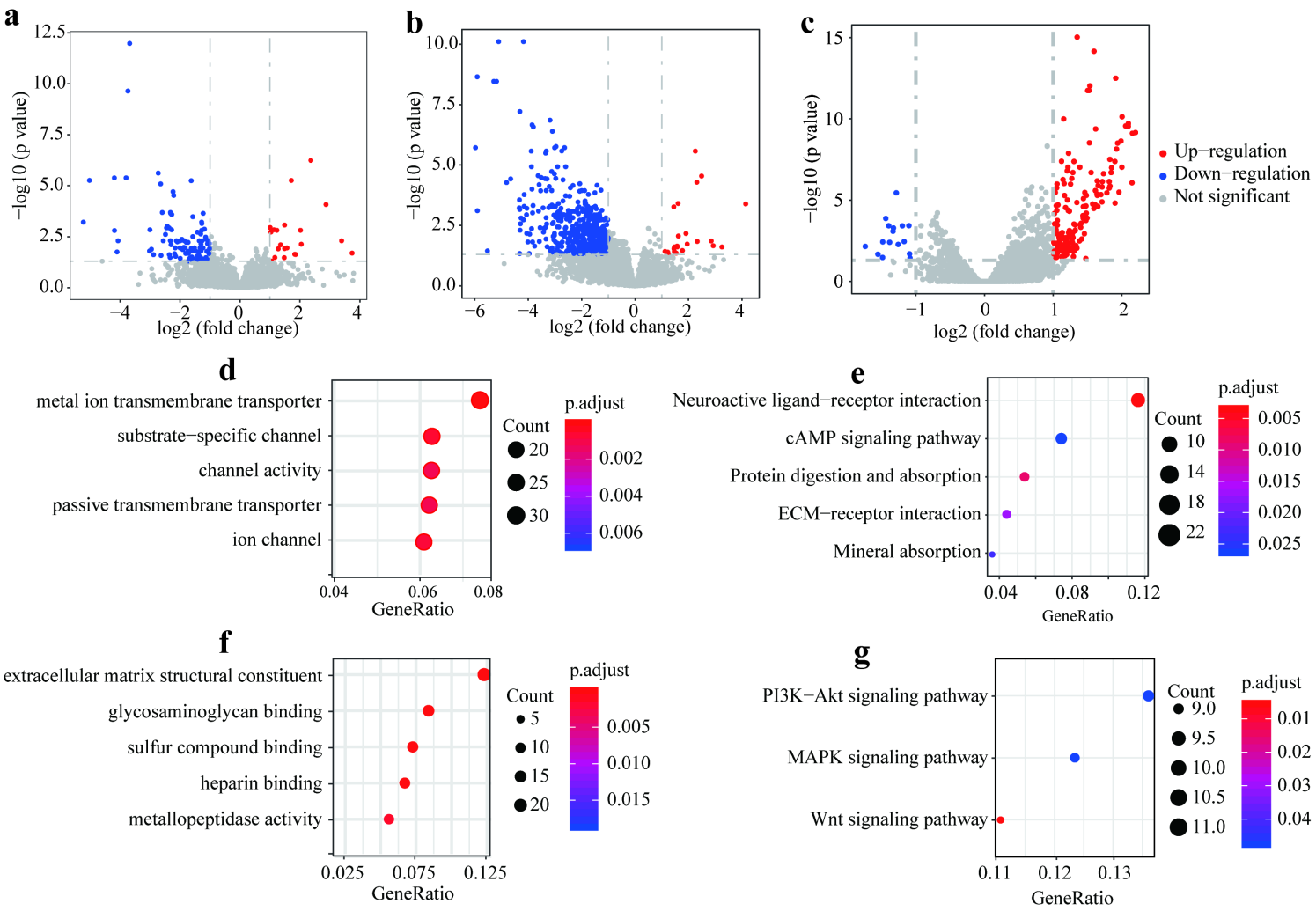

Figure 5

Volcanic map and GO and kEGG annotation map of differential genes. a mRNAsi-high samples were down-regulated by 110 genes and 26 genes were up-regulated compared to mRNAsi-low samples in TCGA. b Compared with mDNAsi-low samples, mDNAsi-high samples 546 genes were down-regulated and 23 genes were up-regulated in TCGA. c mRNAsi-high samples down-regulated 16 genes and up-regulated 174 genes compared to mRNAsi-low samples in ICGC. d Top 5 of GO annotation map of the 569 mDNAsi differential genes in TCGA. e KEGG annotation map of mDNAsi differential genes in TCGA. f GO annotation map of the 569 mRNAsi differential genes in ICGC. g KEGG annotation map of mRNAsi differential genes in ICGC.

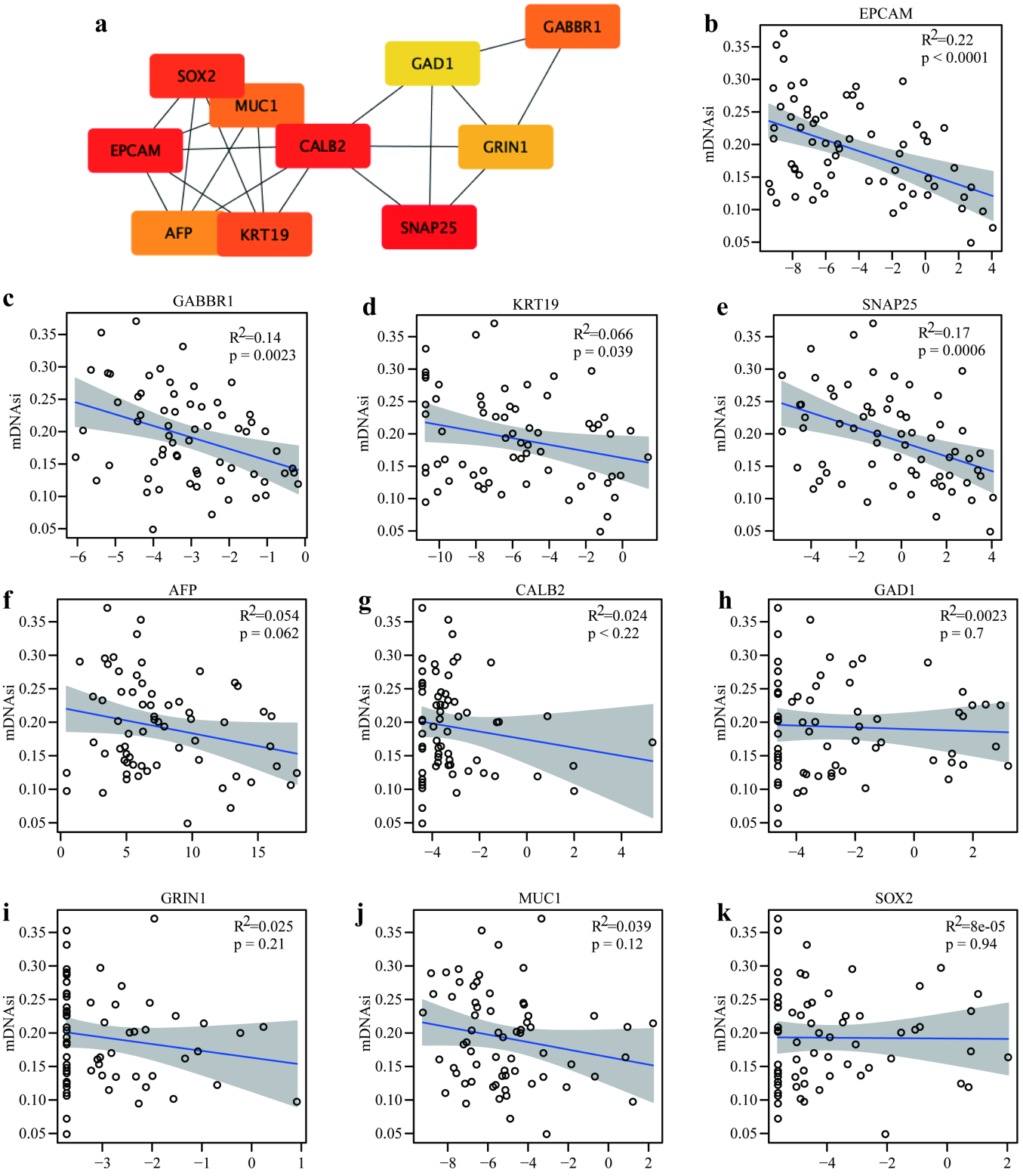

Figure 6

PPI analysis and correlation between Hub genes and mDNAsi. a Protein interaction between 10 hub genes (the darker the color, the higher the score, the more important it is). b EPCAM is significantly negatively correlated with mDNAsi. c GABBR1 is significantly negatively correlated with mDNAsi. d KRT19 is significantly negatively correlated with mDNAsi. e SNAP25 is significantly negatively correlated with mDNAsi. f AFP has no significantly relation with mDNAsi. g CALB2 has no significantly relation with

mDNAsi. h GAD1 has no significantly relation with mDNAsi. i GRIN1 has no significantly relation with mDNAsi. j MUC1 has no significantly relation with mDNAsi. k SOX2 has no significantly relation with mDNAsi.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigureS1.tif
- FigureS2.tif