An immune-related E2F gene expression signature provides prognostic information in patients with hepatocellular carcinoma

Lu Wang,[1] Xiangdong Liu,[1] and Zhongsheng Sun,[2,3]

[1] Key Laboratory of Developmental Genes and Human Diseases, Institute of Life Sciences, Southeast University, 2 Sipailou Road, Nanjing, Jiangsu, 210096, China
[2] Beijing Institutes of Life Science, Chinese Academy of Sciences, No. 1 West Beichen Road, Chaoyang District, Beijing, 100101, China
[3] Institute of Genomic Medicine, Wenzhou Medical University, University Town, Chashan, Wenzhou, Zhejiang, 325035, China

Correspondence should be addressed to Xiangdong Liu; xiangdongliu@seu.edu.cn and Zhongsheng Sun; sunzs@biols.ac.cn

Abstract

Background. Hepatocellular carcinoma (HCC) is a fatal form of liver cancer. Through RB-E2F pathway or other mechanisms, E2F transcription factors impact tumorigenesis and progression of multiple cancers including HCC. We aim to develop a prognostic signature based on the expression of E2Fs for HCC patients. *Methods.* RNA-seq and clinical data from TCGA database were used to identify a prognostic signature based on the expression of E2Fs. Another cohort from GEO database was utilized as external validation. The risk score combined with common clinical factors were employed to establish a nomogram as a clinically-accessible tool. Tumor immune microenvironment influenced by the signature was analyzed and a competing endogenous network of the component genes has been constructed. *Results.* A prognostic signature dependent on the expression of two E2Fs was developed and validated, with the risk score successfully stratifying HCC patients into different risk groups. A nomogram was provided for clinical application. Patients in the high-risk group exhibited significantly higher fractions of immunosuppressive cells than low-risk patients. Moreover, we revealed a mRNA-miRNA-circRNA competing endogenous network that potentially played a pivotal role in the prognosis and progression of HCC. *Conclusion*. This study offered a novel gene signature that could robustly identify HCC patients with poor survival and highlighted potential predictive biomarkers for therapeutic interventions.

1. Introduction

Liver cancer is one of the leading causes of cancer-related mortality worldwide, and its incidence is still growing rapidly although extensive efforts have been made to understand the molecular architecture and prevent morbidity [1]. It has been reported that there were approximately 841,080 newly diagnosed liver cancer cases and 781,631 liver cancer-related death in 2018 [2]. Hepatocellular carcinoma (HCC), as the most prevalent type of primary liver cancer, accounts for up to 90% of all cases [3]. HCC typically develops on a background of chronic liver inflammation or cirrhosis, with hepatitis B virus infection being the most paramount risk factor for its development [4]. Growing evidence suggests the incremental contribution of non-alcoholic steatohepatitis in the context of metabolic syndrome or diabetes mellitus as an etiology of HCC [5]. Despite great improvements in treatment modalities and earlier diagnosis, the long-term prognosis of HCC patients after surgery remains dismal. Current TNM staging system and other traditional clinical models are useful in providing prognostic assessment, however far from being ideally accurate in the grading of patients for surgical procedures.

Unlike cells in other organs or tissues, the regeneration of which is elicited by stem cells, hepatocytes are characterized by an impressive ability of maintaining remarkable proliferative potential to compensate for cell loss or in response to hyperplasia [6]. Members of the E2F transcription factor family are important in regulation of hepatocyte proliferation, differentiation, apoptosis, and tumorigenesis in collaboration with cyclin-dependent kinases (CDKs) and the RB family of proteins [7-11]. As previously demonstrated, dysfunction of the CDK-RB-E2F axis would lead to aberrant cell cycle and inhibition of apoptosis, and thus intimately related to pathological processes of HCC [8,12,13]. The E2F transcription factors have traditionally been classified into activators and repressors, but this classification is now been challenged as a gross oversimplification of the complexity of this family neglecting the existence of overlapping functions and interchangeable use between these E2Fs [14-16]. Moreover, functions of dysregulated E2F activities were shown to be exquisitely dependent on cell type and physiological contexts [17-20]. Aberrant expression of the E2Fs have been implicated in several human malignancies [21-24]. Elevated E2F target expression have been reported to be associated with poor survival and predictive of prognosis in different tumor types [25-27].

In this study, we performed Cox proportional hazards regression analysis to develop a gene expression prognostic signature of the E2F family members based on the RNA-seq data from the TCGA database and validated its predictive performance with an external validation cohort from the GEO database. Two E2F transcription factors, E2F2 and E2F5, were finally retained in the prognostic signature. Based on the risk prediction of this signature, we revealed that the HCC patients classified into the high-risk group had significantly lower survival rates than the patients in low-risk group. To further explore and compare the tumor infiltration patterns between patients of different risk groups, the proportions of diverse tumor-infiltrating immune cell types has been estimated using CIBERSORT algorithm. We also constructed a mRNA-miRNA-circRNA competing endogenous network for E2F2 and E2F5 seeking to provide insights into the regulatory mechanisms of the component genes. In

addition, a nomogram integrating the prognostic signature with common clinicopathological parameters was established to enhance the predictive capacity and to provide clinicians with a quantitative and easily accessible tool.

2. Materials and Methods

2.1. Retrieval of RNA-sequencing Data. The level III transcriptome profiles including mRNA-seq data and miRNA-seq data of 369 patients with HCC and the corresponding clinical information were downloaded from the TCGA-LIHC dataset. The clinical parameters of age, gender, family history of cancer, pathologic stage, histologic grade, Ishak fibrosis score, Child–Pugh grade, vascular invasion, alpha fetoprotein outcome and residual tumor were evaluated in this study. After removing samples with incomplete survival information, 339 HCC patients were retained for subsequent analysis. The patients with HCC were randomly assigned to a training set or a validation set at a ratio of 1:1 (n = 171 for training set and n = 168 for validation set).

2.2. Construction of the Prognostic Signature. To establish a clinically translatable prognostic-related gene signature, we initially performed univariate Cox proportional hazards regression analysis on the genes of the E2F family using expression data from the TCGA training set. Genes with a P value less than 0.05 were considered as significant genes and kept for subsequent stepwise multivariate Cox analysis. Finally, the prognostic-related gene signature was built and the risk score derived from the gene signature was calculated as the sum of the expression level of each E2F gene multiplied by the corresponding regression coefficients obtained from the multivariate Cox analysis. The median risk score of the training cohort was chosen as the cutoff point to stratify the patients into the high-risk and low-risk groups for the whole set of TCGA HCC patients. Kaplan Meier survival analysis and the log-rank test were utilized to examine the predictive performance for this signature.

2.3. External Validation of the Prognostic Signature. The GSE14520 dataset consisting of 222 HCC patients with complete clinical information was retrieved from the Gene Expression Omnibus database. The risk score for each patient in the GSE14520 validation dataset was calculated based on the prognostic signature and stratified into high-risk and low-risk groups by the median risk score derived from the TCGA training cohort. Kaplan Meier survival analysis was carried out to assess the predictive ability.

2.4. Building a Quantitative Nomogram Based on the Prognostic Signature. To individualize the predicted results on the 1-year, 3-year and 5-year survival probability for HCC patients and further improve the predictive accuracy of our prognostic signature, we established a nomogram as a quantitative tool based on the risk score derived from the prognostic signature combined with common clinical characteristics. Kaplan Meier survival analysis was performed and the C-index was calculated to determine the discrimination capacity of the nomogram. Calibration plots were used to

compare and display the agreement between the predicted results and the observed outcomes with regard to the 1-year, 3-year and 5-year overall survival for the HCC patients from the TCGA cohort.

2.5. Estimation of the Tumor-Infiltrating Immune Cells. To demonstrate the differences in the proportions of diverse tumor-infiltrating immune cells between the high-risk and low-risk samples, we utilized the CIBERSORT algorithm in combination with the LM22 signature matrix consisting of 547 genes to derive the immune infiltration status for 22 mature human hematopoietic populations in the HCC patients. The permutation test of 1,000 iterations was adopted to obtain the result of statistical significance followed by quantile normalization.

2.6. Inference of Regulatory Interaction Relationships for the Component Genes in the Prognostic Signature. The mRNA-miRNA interaction relationships for the component genes from the prognostic signature were acquired from the miRTarBase compiled in the NetworkAnalyst platform. The circRNA-miRNA interaction data were accessed from StarBase. The miRNA binding sites located in the 3'-UTR of the component genes were predicted using the TargetScan algorithm.

2.7. Expression and survival analysis for E2F2 and E2F5. The UALCAN (http://ualcan.path.uab.edu) web server was utilized to investigate the expression levels of E2F2 and E2F5 between tumor and normal tissues, as well as across different clinical condition groups. Survival analysis for the relationship between gene expression level and overall survival rate of HCC patients was carried out using GEPIA (http://gepia.cancer-pku.cn/detail.php). HCC patients were classified into two different groups according to the median expression level of the respective gene.

3. Results

3.1. Construction and Evaluation of A Two-E2F Signature Associated with Poor Prognosis in HCC Patients. Based on the expression data and clinical information from the TCGA HCC cohort, we applied Cox proportional hazards regression analysis on the E2F family members. 339 HCC patients from the TCGA database were randomly divided into a training set (n=171) and a validation set (n=168). After conducting a stepwise multivariate Cox analysis on the training cohort, E2F2 and E2F5 were identified to be significantly correlated with poor survival of the HCC patients (Table). Within the prognostic signature, a risk score was calculated for each patient according to the expression level of E2F2 and E2F5. Based on the median risk score for the training cohort, the HCC patients were categorized into different risk groups. We observed that patients in the low-risk group exhibited remarkably longer overall survival (OS) compared to high-risk patients in the TCGA training cohort ($p = 2.229e{-}04$, Figure 1(a)) as well as the validation cohort ($p = 1.638e{-}02$, Figure 1(b)).

The time-dependent receiver operating characteristic (ROC) curves were applied to evaluated the predicting capacity of the two-E2F prognostic signature. Area under the curve (AUC) yielded by the prognostic signature at 1-year, 3-year and 5-year survival prediction was 0.749, 0.683 and 0.698, respectively, for the TCGA training cohort (Figure S1(a)). We further confirmed the prediction efficiency for the two-E2F signature in the validation cohort and obtained the AUCs for 1-year, 3-year and 5-year OS rates of 0.661, 0.641 and 0.659, respectively (Figure S1(b)). In addition, it was reassuring to note that the two-E2F prognostic signature was able to robustly stratify patients into high- and low-risk groups for OS rates in an external validation cohort (GSE14520, n=222, Figure 1(c)). These findings demonstrated the effectiveness of the two-E2F signature in predicting unfavorable prognosis for HCC patients.

3.2. The Prognostic Risk Score Correlated with Pathological Grade and Tumor Progression. To further evaluate the potential of the two-E2F prognostic signature, we investigated the relationship between the risk level and the potential clinical outcomes using the whole set of 339 TCGA patients. It has been revealed that there exists a significant difference in several clinicopathological factors such as TNM stage, histologic grade, Ishak score, Child–Pugh grade, vascular invasion, and the serum alpha fetoprotein between the two risk groups. However, the risk score was not significantly associated with age, gender and family history of cancer (Table 1).

3.3. The Two-E2F Signature Showed A Predictive Role Independent of Conventional Clinicopathological Parameters. We next performed Cox regression analysis on the two-E2F signature-derived risk score incorporating the commonly used clinicopathologic parameters. Multivariate Cox analysis including only the features that were statistically significant in the univariate analysis demonstrated that the risk score obtained from the two-E2F signature presented as an independent prognostic predictor for HCC patients (Table 2).

3.4. Estimation of immune Infiltration Levels for Patients from Different Risk Groups. In order to elucidate the differences in immune microenvironment composition of samples between high-risk and low-risk groups, we applied CIBERSORT algorithm to infer the fraction of different types of immune cells using the TCGA cohort expression data. The immune cell distribution was calculated based on a matrix of 22 immune cell type signatures (LM22) and compared between high-risk and low-risk HCC patients (Figure 2(a)). The proportions of tumor-infiltrating macrophages M0 and T cells regulatory (Tregs) were significantly higher in the high-risk HCC patients than those in the low-risk group, while the low-risk patients had markedly higher proportions of T cells CD4 memory resting than the high-risk HCC patients (Figure 2(b)). These results highlighted a potential mechanistic link between heterogeneous immune infiltration status across different HCC patients and variable clinical manifestations.

3.5. Establishment of A Nomogram Based on the Two-E2F Prognostic Signature. With the goal of providing a clinically adaptable approach to quantifying the risk assessment

for HCC patients, we developed a nomogram integrated the two-E2F signature and the commonly used clinicopathologic risk factors with the whole set of TCGA samples (Figure 3(a)). We calculated the C-statistic discriminatory index (C-index) and the calibration plot to assess the prognostic performance for the nomogram. As shown in the calibration plot, there is excellent agreement between the predicted outcomes and the observed 1-year, 3-year and 5-year OS (Figure 3(b)). And the C-index for the nomogram was 0.712 (95% CI: 0.683 - 0.742). In addition, Kaplan Meier analysis based on the median risk score derived from the nomogram integrating the two-E2F signature and the clinicopathologic features was also carried out in order to measure the discriminatory power of this nomogram. The patients being labeled as high-risk exhibited significantly poorer prognosis than the low-risk patients (Figure 3(c)). These findings demonstrated the validity of our two-E2F signature in combination with the commonly used clinicopathologic factors in predicting survival outcomes for HCC patients.

3.6. Further Excavation of E2F2 and E2F5. Given the above lines of evidence supporting the role of the two-E2F signature in predicting prognosis for HCC patients, we then inspected the two component genes in the prognostic signature (E2F2 and E2F5) to provide further insights into the mechanisms by which the E2Fs may exert effects in the pathogenesis of HCC. Using expression data from the TCGA cohort, prominent upregulation of E2F2 and E2F5 in tumor samples in comparison to normal controls have been revealed (Figures 4(a) and 4(b)). Moreover, expression levels of E2F2 and E2F5 were both significantly elevated in patients with advanced tumor stage and higher histologic grade (Figures 4(a) and 4(b)). In addition, since TP53 is one of the most frequently mutated genes in patients with HCC [28], we divided the TCGA cohort into different groups based on TP53 mutation status and evaluated the expression levels of E2F2 and E2F5 between TP53 mutation group and TP53 wild-type group. It has been shown that expression levels of both E2F2 and E2F5 were significantly associated with the mutation status of TP53 and elevated in patients mutated in TP53 (Figures 4(a) and 4(b)). As expected, we also found that higher expression levels of E2F2 and E2F5 were significantly associated with poorer survival of HCC patients (Figures 4(c) and 4(d)).

In an attempt to explore the regulatory interaction relationships involved in the expression of E2F2 and E2F5, we constructed an miRNA-mRNA network to identify the key upstream miRNA regulators for these two genes according to the experimentally validated miRNA-target interactions deposited in miRTarBase (Figure S2). Dysregulation of several miRNAs has been observed including miR-125b-5p and miR-22-3p targeting the E2F2 gene and let-7c-5p that were predicted to regulate E2F5 (Figures 5(a)-5(c)). Furthermore, expression changes of these three miRNAs exhibited significant negative correlations with the respective target E2Fs (Figures 5(d)-5(f)). Kaplan-Meier analysis revealed that expression level of miR-125b-5p, miR-22-3p and let-7c-5p were all significantly associated with lower survival rates of HCC patients (Figures 5(g)-5(i)).

Intriguingly, according to the predictions from the StarBase database, we found that circ-C19orf54 acted as miRNA sponge for both miR-125b-5p and let-7c-5p, which

targeted E2F2 and E2F5, respectively (Figures 6(a)-6(c)). We analyzed the expression of C19orf54 and revealed that the expression level of C19orf54 was dramatically up-regulated in HCC patients in comparison to normal controls from the TCGA cohort (Figure 6(d)). Moreover, significant positive correlation between the expression of circRNA C19orf54 with the expression level of E2F2 and E2F5, respectively, has also been observed using the TCGA-HCC dataset (Figures 6(e) and 6(f)). These results demonstrated the existence of a potential competing endogenous network formed by circ-C19orf54, miR-125b-5p, let-7c-5p, E2F2 and E2F5.

4. Discussion

During the progression of cell cycle, the regulation of cell proliferation involves a cascade of molecular interactions associated with genome stability, DNA repair and chromatin regulation [29]. Mechanistic studies have exposed a pivotal role for the CDK-RB-E2F axis in the maintenance of genomic integrity and cell homeostasis, as well as responding to DNA damage and stress [30-33]. The CDK inhibitors have long been regarded as promising targets for therapeutic intervention in cancer, mediating the phosphorylation and activation of RB proteins and thus enabling the release of E2F transcription factors that regulate a plethora of genes required for DNA synthesis and chromosome stability during cell cycle progression [34-37]. An unequivocal consensus has been reached on the central role of the CDK-RB-E2F pathway in cancer [38,39]. The E2Fs represent a crucial group of transcriptional regulators that are involved in cell cycle control and consists of nine different gene products encoded by eight chromosomal loci. Based on in vitro structural studies, the E2F family members have been artificially subdivided into three categories according to the sequence homology and protein function [40,41], although challenged by in vivo studies as being a gross oversimplification [16,42]. While the precise mechanisms remain unclear and warrant further research, dysregulated functions of the E2F family members and/or the target genes associated with apoptosis, metabolism or angiogenesis have been linked to poor outcomes in cancers [10,25-27]. In this report, we established a prognostic signature based on the expression profiles of the E2F transcription factors that could stratify patients with HCC into different risk groups. Patients classified into the high-risk group exhibited significantly more unfavorable prognosis than the low-risk patients. As the component genes in the signature, E2F2 and E2F5 may act in combination to coordinate a series of cellular processes involved in cancer aggressiveness.

E2F2 is traditionally classified as an activator and it has been reported that the E2F2 gene played a pivotal role in liver regeneration, the absence of which resulted in deregulation of S-phase entry and inhibition of hepatocytes proliferation in mice [43]. Previous reports have demonstrated the contributory role of E2F2 upregulation in the development of breast cancer, owing to the enrichment of its target genes related to angiogenesis and extracellular matrix remodeling [44]. It has been revealed that E2F2 also possess pro-apoptotic functions in the scenario of DNA damage [45]. In HCC, there have been evidence suggesting that down-regulation of E2F2 through a miRNA-mediated machinery could result in the inhibition of hepatocyte proliferation and

prevention of cell progression [46]. Rather than being a traditionally acknowledged activator as E2F2, E2F5 has been grouped into the repressor group based on previous in vitro studies [42,47]. It has been demonstrated that the proliferation and migration of cells in breast cancer can be inhibited by decreasing the transcriptional output of E2F5 through a miRNA-dependent mechanism [48]. In addition, E2F5 was also found to be substantially overexpressed in glioblastoma and prostate cancer, under both conditions involving miRNAs as regulators [49,50]. Furthermore, several studies have sought to explore the relationship between E2F5 and HCC, and the results obtained revealed that the dysregulation of E2F5 expression was associated with the process of tumor development [51,52].

In addition to the expression dysregulation of the constituent gens revealed by using RNA-seq data publicly available from TCGA, we demonstrated that the two-E2F prognostic signature served as an independent risk predictor for HCC patients after adjusting for a set of conventional clinical parameters using multivariate Cox analysis. Besides, both component genes of the prognostic signature were proved to be individually associated with shortened survival. To further examine the efficacy of the risk score derived from the prognostic signature, we analyzed an external validation cohort from the GEO dataset and revealed that the risk score obtained from the two-E2F prognostic signature can effectively stratify the HCC patients into high-risk and low-risk groups for overall survival. Taken together, these observations supported the role of the two-E2F signature as a novel and robust predictor for survival outcomes in patients with HCC. Moreover, to provide a clinically applicable tool for clinicians, we constructed a nomogram integrating the risk score derived from the prognostic signature with the commonly used clinical characteristics for a more targeted therapy and tailed treatment in individual patients. This nomogram offering a personalized survival rates estimate system for patients with HCC could be implemented as an efficient predictive tool in clinical settings.

During the process of tumor cell proliferation and tumor growth, a disturbed immune microenvironment could be involved in promoting tumoral immune escape [53]. Here. we demonstrated the differences in infiltration of immune cells between patients in the high-risk and low-risk groups. Remarkably, patients in the high-risk group showed higher proportions of immunosuppressive cells such as T cells follicular helper, T cells regulatory and macrophages M0 than the low-risk patients, suggesting that the observed decreased survival rates in high-risk patients could result from the enhanced immunosuppressive environment induced by cancer cells.

In order to gain further insights into the regulatory mechanisms of E2F2 and E2F5, we identified the miRNAs targeting these two genes using the experimentally validated interactions deposited in the miRTarBase and focused on miR-125b-5p, miR-22-3p and let-7c-5p that showed significantly decreased expression in tumors relative to normal samples. Moreover, strong negative correlations have been observed between these three miRNAs and their respective targets. Reduced expression of miR-125b-5p has been observed in HCC and related to inhibition of cell proliferation, migration, and invasion according to previous studies [54-56]. Downregulation of miR-125b-5p has been associated with poor prognosis of HCC patients, which is consistent with our

current findings [54,55]. Several studies have sought to uncover the role of miR-22-3p in the tumorigenesis of HCC and revealed that dysfunction of miR-22-3p was involved in cell proliferation, cellular senescence and the tumor-suppressing phenotypes in HCC [57]. Notably, there is also evidence indicating that let-7c-5p and its target genes are collectively responsible for tumor growth, apoptosis and metastasis in HCC [58]. Moreover, we found that the circular RNA C19orf54 was regulated by miR-125b-5p as well as let-7c-5p, both of which have been shown tumor suppressing properties, suggesting that the expression of E2F2 and E2F5 could be mediated by the competing endogenous RNA mechanism at the posttranscriptional level with circRNA C19orf54 acting as an miRNA sponge. Remarkably, we observed that the expression level of C19orf54 was significantly upregulated in HCC compared with normal controls and positively correlated with E2F2 and E2F5, which further supported the proposed mechanism.

However, in this study, lack of complete clinicopathological records and information about treatment regimens resulted in loss of samples and could contribute to decreased statistical power. In addition, prospective studies using independent cohorts should be carried out in the future to assess and improve the quality of these findings.

In conclusion, we have identified and validated a prognostic signature that requires only two E2F transcription factors. The clinically accessible quantitative nomogram based on the two-E2F signature allows robust stratification of risk subtypes in HCC. Furthermore, we described the landscape of immune infiltration for HCC patients, highlighting the important roles of tumor microenvironment adaptations and immune cell heterogeneity. Notably, circC19orf54, miR-125b-5p, let-7c-5p, E2F2 and E2F5 can form a competing endogenous network that are associated with the prognosis of HCC patients and have great potential to serve as novel predictive biomarkers and therapeutic targets for clinical interventions.

5. Conclusions

In this study, we identified and validated a prognostic signature based on expressions of two E2F transcription factors for patients with HCC and built a clinically applicable nomogram as a quantitative tool for clinician. Moreover, it has been revealed that patients in the high-risk group exhibited significantly higher fractions of immunosuppressive cells such as Treg cells and tumor-associated macrophages, suggesting the feasibility of improving clinical efficacy for cancer immunotherapy on the high-risk HCC patients based on the heterogeneity of immune infiltration. After investigating the interaction relationships between E2F2, E2F5, hsa-miR-125b-5p, hsa-let-7c-5p and circRNA C19orf54, a competing endogenous RNA model was finally proposed and these molecules have the potential to be promising predictive biomarkers for therapeutic intervention to improve HCC prognosis and personalized cancer management.

Data Availability

The source data used to support the findings of this study were derived from the public repositories, as indicated in the section of "Materials and Methods" of the manuscript. And all data that support the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Authors' Contributions

L.W., X.D.L. and Z.S.S. conceived the study. L.W., X.D.L. and Z.S.S. designed the study. L.W. analyzed the data. L.W. and X.D.L. interpreted the results and produced the original draft. Z.S.S. edited the manuscript. X.D.L. and Z.S.S. supervised the study. All authors have read and agreed to the published version of the manuscript.

References

1. Villanueva, A. Hepatocellular Carcinoma. *N Engl J Med* **2019**, *380*, 1450-1462, doi:10.1056/NEJMra1713263.
2. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **2018**, *68*, 394-424, doi:10.3322/caac.21492.
3. European Association for the Study of the Liver. Electronic address, e.e.e.; European Association for the Study of the, L. EASL Clinical Practice Guidelines: Management of hepatocellular carcinoma. *Journal of hepatology* **2018**, *69*, 182-236, doi:10.1016/j.jhep.2018.03.019.
4. Ganne-Carrie, N.; Nahon, P. Hepatocellular carcinoma in the setting of alcohol-related liver disease. *Journal of hepatology* **2019**, *70*, 284-293, doi:10.1016/j.jhep.2018.10.008.
5. Estes, C.; Razavi, H.; Loomba, R.; Younossi, Z.; Sanyal, A.J. Modeling the epidemic of nonalcoholic fatty liver disease demonstrates an exponential increase in burden of disease. *Hepatology* **2018**, *67*, 123-133, doi:10.1002/hep.29466.
6. Malato, Y.; Naqvi, S.; Schurmann, N.; Ng, R.; Wang, B.; Zape, J.; Kay, M.A.; Grimm, D.; Willenbring, H. Fate tracing of mature hepatocytes in mouse liver homeostasis and regeneration. *J Clin Invest* **2011**, *121*, 4850-4860, doi:10.1172/JCI59261.
7. Hassan, M.; Ghozlan, H.; Abdel-Kader, O. Activation of RB/E2F signaling pathway is required for the modulation of hepatitis C virus core protein-induced cell growth in liver and non-liver cells. *Cell Signal* **2004**, *16*, 1375-1385, doi:10.1016/j.cellsig.2004.04.005.
8. Mayhew, C.N.; Carter, S.L.; Fox, S.R.; Sexton, C.R.; Reed, C.A.; Srinivasan, S.V.; Liu, X.; Wikenheiser-Brokamp, K.; Boivin, G.P.; Lee, J.S., et al. RB loss abrogates cell cycle control and genome integrity to promote liver tumorigenesis. *Gastroenterology* **2007**, *133*, 976-984, doi:10.1053/j.gastro.2007.06.025.
9. Kent, L.N.; Leone, G. The broken cycle: E2F dysfunction in cancer. *Nat Rev Cancer* **2019**, *19*, 326-338, doi:10.1038/s41568-019-0143-7.

10. Chen, H.Z.; Tsai, S.Y.; Leone, G. Emerging roles of E2Fs in cancer: an exit from cell cycle control. *Nat Rev Cancer* **2009**, *9*, 785-797, doi:10.1038/nrc2696.
11. Attwooll, C.; Lazzerini Denchi, E.; Helin, K. The E2F family: specific functions and overlapping interests. *EMBO J* **2004**, *23*, 4709-4716, doi:10.1038/sj.emboj.7600481.
12. Rowland, B.D.; Bernards, R. Re-evaluating cell-cycle regulation by E2Fs. *Cell* **2006**, *127*, 871-874, doi:10.1016/j.cell.2006.11.019.
13. Hernando, E.; Nahle, Z.; Juan, G.; Diaz-Rodriguez, E.; Alaminos, M.; Hemann, M.; Michel, L.; Mittal, V.; Gerald, W.; Benezra, R., et al. Rb inactivation promotes genomic instability by uncoupling cell cycle progression from mitotic control. *Nature* **2004**, *430*, 797-802, doi:10.1038/nature02820.
14. Chong, J.L.; Wenzel, P.L.; Saenz-Robles, M.T.; Nair, V.; Ferrey, A.; Hagan, J.P.; Gomez, Y.M.; Sharma, N.; Chen, H.Z.; Ouseph, M., et al. E2f1-3 switch from activators in progenitor cells to repressors in differentiating cells. *Nature* **2009**, *462*, 930-934, doi:10.1038/nature08677.
15. Martinez, L.A.; Goluszko, E.; Chen, H.Z.; Leone, G.; Post, S.; Lozano, G.; Chen, Z.; Chauchereau, A. E2F3 is a mediator of DNA damage-induced apoptosis. *Mol Cell Biol* **2010**, *30*, 524-536, doi:10.1128/MCB.00938-09.
16. Chen, Y.L.; Uen, Y.H.; Li, C.F.; Horng, K.C.; Chen, L.R.; Wu, W.R.; Tseng, H.Y.; Huang, H.Y.; Wu, L.C.; Shiue, Y.L. The E2F transcription factor 1 transactives stathmin 1 in hepatocellular carcinoma. *Ann Surg Oncol* **2013**, *20*, 4041-4054, doi:10.1245/s10434-012-2519-8.
17. Kontaki, H.; Talianidis, I. Lysine methylation regulates E2F1-induced cell death. *Mol Cell* **2010**, *39*, 152-160, doi:10.1016/j.molcel.2010.06.006.
18. Lin, W.C.; Lin, F.T.; Nevins, J.R. Selective induction of E2F1 in response to DNA damage, mediated by ATM-dependent phosphorylation. *Genes Dev* **2001**, *15*, 1833-1844.
19. Stevens, C.; Smith, L.; La Thangue, N.B. Chk2 activates E2F-1 in response to DNA damage. *Nat Cell Biol* **2003**, *5*, 401-409, doi:10.1038/ncb974.
20. Engelmann, D.; Putzer, B.M. Translating DNA damage into cancer cell death-A roadmap for E2F1 apoptotic signalling and opportunities for new drug combinations to overcome chemoresistance. *Drug Resist Updat* **2010**, *13*, 119-131, doi:10.1016/j.drup.2010.06.001.
21. Santos, M.; Martinez-Fernandez, M.; Duenas, M.; Garcia-Escudero, R.; Alfaya, B.; Villacampa, F.; Saiz-Ladera, C.; Costa, C.; Oteo, M.; Duarte, J., et al. In vivo disruption of an Rb-E2F-Ezh2 signaling loop causes bladder cancer. *Cancer Res* **2014**, *74*, 6565-6577, doi:10.1158/0008-5472.CAN-14-1218.
22. Rennhack, J.; Andrechek, E. Conserved E2F mediated metastasis in mouse models of breast cancer and HER2 positive patients. *Oncoscience* **2015**, *2*, 867-871, doi:10.18632/oncoscience.259.
23. Shackney, S.E.; Chowdhury, S.A.; Schwartz, R. A Novel Subset of Human Tumors That Simultaneously Overexpress Multiple E2F-responsive Genes Found in Breast, Ovarian, and Prostate Cancers. *Cancer Inform* **2014**, *13*, 89-100, doi:10.4137/CIN.S14062.
24. Feliciano, A.; Garcia-Mayea, Y.; Jubierre, L.; Mir, C.; Hummel, M.; Castellvi, J.; Hernandez-Losa, J.; Paciucci, R.; Sansano, I.; Sun, Y., et al. miR-99a reveals two novel oncogenic proteins E2F2 and EMR2 and represses stemness in lung cancer. *Cell Death Dis* **2017**, *8*, e3141, doi:10.1038/cddis.2017.544.
25. Lan, W.; Bian, B.; Xia, Y.; Dou, S.; Gayet, O.; Bigonnet, M.; Santofimia-Castano, P.; Cong, M.; Peng, L.; Dusetti, N., et al. E2F signature is predictive for the pancreatic adenocarcinoma clinical outcome and sensitivity to E2F inhibitors, but not for the response to cytotoxic-based treatments. *Sci Rep* **2018**, *8*, 8330, doi:10.1038/s41598-018-26613-z.

26. Kent, L.N.; Rakijas, J.B.; Pandit, S.K.; Westendorp, B.; Chen, H.Z.; Huntington, J.T.; Tang, X.; Bae, S.; Srivastava, A.; Senapati, S., et al. E2f8 mediates tumor suppression in postnatal liver development. *J Clin Invest* **2016**, *126*, 2955-2969, doi:10.1172/JCI85506.
27. Kent, L.N.; Bae, S.; Tsai, S.Y.; Tang, X.; Srivastava, A.; Koivisto, C.; Martin, C.K.; Ridolfi, E.; Miller, G.C.; Zorko, S.M., et al. Dosage-dependent copy number gains in E2f1 and E2f3 drive hepatocellular carcinoma. *J Clin Invest* **2017**, *127*, 830-842, doi:10.1172/JCI87583.
28. Schulze, K.; Imbeaud, S.; Letouze, E.; Alexandrov, L.B.; Calderaro, J.; Rebouissou, S.; Couchy, G.; Meiller, C.; Shinde, J.; Soysouvanh, F., et al. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nature genetics* **2015**, *47*, 505-511, doi:10.1038/ng.3252.
29. Otto, T.; Sicinski, P. Cell cycle proteins as promising targets in cancer therapy. *Nat Rev Cancer* **2017**, *17*, 93-115, doi:10.1038/nrc.2016.138.
30. Ohtani, N.; Brennan, P.; Gaubatz, S.; Sanij, E.; Hertzog, P.; Wolvetang, E.; Ghysdael, J.; Rowe, M.; Hara, E. Epstein-Barr virus LMP1 blocks p16INK4a-RB pathway by promoting nuclear export of E2F4/5. *J Cell Biol* **2003**, *162*, 173-183, doi:10.1083/jcb.200302085.
31. Wetmore, C.; Boyett, J.; Li, S.; Lin, T.; Bendel, A.; Gajjar, A.; Orr, B.A. Alisertib is active as single agent in recurrent atypical teratoid rhabdoid tumors in 4 children. *Neuro Oncol* **2015**, *17*, 882-888, doi:10.1093/neuonc/nov017.
32. Kastan, M.B.; Bartek, J. Cell-cycle checkpoints and cancer. *Nature* **2004**, *432*, 316-323, doi:10.1038/nature03097.
33. Bartek, J.; Lukas, C.; Lukas, J. Checking on DNA damage in S phase. *Nat Rev Mol Cell Biol* **2004**, *5*, 792-804, doi:10.1038/nrm1493.
34. Hutcheson, J.; Witkiewicz, A.K.; Knudsen, E.S. The RB tumor suppressor at the intersection of proliferation and immunity: relevance to disease immune evasion and immunotherapy. *Cell Cycle* **2015**, *14*, 3812-3819, doi:10.1080/15384101.2015.1010922.
35. Malumbres, M.; Barbacid, M. Mammalian cyclin-dependent kinases. *Trends Biochem Sci* **2005**, *30*, 630-641, doi:10.1016/j.tibs.2005.09.005.
36. Sherr, C.J.; Roberts, J.M. Living with or without cyclins and cyclin-dependent kinases. *Genes Dev* **2004**, *18*, 2699-2711, doi:10.1101/gad.1256504.
37. Hydbring, P.; Malumbres, M.; Sicinski, P. Non-canonical functions of cell cycle cyclins and cyclin-dependent kinases. *Nat Rev Mol Cell Biol* **2016**, *17*, 280-292, doi:10.1038/nrm.2016.27.
38. Sherr, C.J. Cancer cell cycles. *Science* **1996**, *274*, 1672-1677, doi:10.1126/science.274.5293.1672.
39. Sherr, C.J.; McCormick, F. The RB and p53 pathways in cancer. *Cancer Cell* **2002**, *2*, 103-112, doi:10.1016/s1535-6108(02)00102-2.
40. Lammens, T.; Li, J.; Leone, G.; De Veylder, L. Atypical E2Fs: new players in the E2F transcription factor family. *Trends Cell Biol* **2009**, *19*, 111-118, doi:10.1016/j.tcb.2009.01.002.
41. Iaquinta, P.J.; Lees, J.A. Life and death decisions by the E2F transcription factors. *Curr Opin Cell Biol* **2007**, *19*, 649-657, doi:10.1016/j.ceb.2007.10.006.
42. Chen, H.Z.; Ouseph, M.M.; Li, J.; Pecot, T.; Chokshi, V.; Kent, L.; Bae, S.; Byrne, M.; Duran, C.; Comstock, G., et al. Canonical and atypical E2Fs regulate the mammalian endocycle. *Nat Cell Biol* **2012**, *14*, 1192-1202, doi:10.1038/ncb2595.
43. Delgado, I.; Fresnedo, O.; Iglesias, A.; Rueda, Y.; Syn, W.K.; Zubiaga, A.M.; Ochoa, B. A role for transcription factor E2F2 in hepatocyte proliferation and timely liver regeneration. *Am J Physiol Gastrointest Liver Physiol* **2011**, *301*, G20-31, doi:10.1152/ajpgi.00481.2010.

44. Hollern, D.P.; Honeysett, J.; Cardiff, R.D.; Andrechek, E.R. The E2F transcription factors regulate tumor development and metastasis in a mouse model of metastatic breast cancer. *Mol Cell Biol* **2014**, *34*, 3229-3243, doi:10.1128/MCB.00737-14.
45. Chen, D.; Chen, Y.; Forrest, D.; Bremner, R. E2f2 induces cone photoreceptor apoptosis independent of E2f1 and E2f3. *Cell Death Differ* **2013**, *20*, 931-940, doi:10.1038/cdd.2013.24.
46. Dong, Y.; Zou, J.; Su, S.; Huang, H.; Deng, Y.; Wang, B.; Li, W. MicroRNA-218 and microRNA-520a inhibit cell proliferation by downregulating E2F2 in hepatocellular carcinoma. *Mol Med Rep* **2015**, *12*, 1016-1022, doi:10.3892/mmr.2015.3516.
47. Tsai, S.Y.; Opavsky, R.; Sharma, N.; Wu, L.; Naidu, S.; Nolan, E.; Feria-Arias, E.; Timmers, C.; Opavska, J.; de Bruin, A., et al. Mouse development with a single E2F activator. *Nature* **2008**, *454*, 1137-1141, doi:10.1038/nature07066.
48. Xu, H.; Fei, D.; Zong, S.; Fan, Z. MicroRNA-154 inhibits growth and invasion of breast cancer cells through targeting E2F5. *Am J Transl Res* **2016**, *8*, 2620-2630.
49. Fang, D.Z.; Wang, Y.P.; Liu, J.; Hui, X.B.; Wang, X.D.; Chen, X.; Liu, D. MicroRNA-129-3p suppresses tumor growth by targeting E2F5 in glioblastoma. *Eur Rev Med Pharmacol Sci* **2018**, *22*, 1044-1050, doi:10.26355/eurrev_201802_14387.
50. Li, S.L.; Sui, Y.; Sun, J.; Jiang, T.Q.; Dong, G. Identification of tumor suppressive role of microRNA-132 and its target gene in tumorigenesis of prostate cancer. *Int J Mol Med* **2018**, *41*, 2429-2433, doi:10.3892/ijmm.2018.3421.
51. Jiang, Y.; Yim, S.H.; Xu, H.D.; Jung, S.H.; Yang, S.Y.; Hu, H.J.; Jung, C.K.; Chung, Y.J. A potential oncogenic role of the commonly observed E2F5 overexpression in hepatocellular carcinoma. *World J Gastroenterol* **2011**, *17*, 470-477, doi:10.3748/wjg.v17.i4.470.
52. Zou, C.; Li, Y.; Cao, Y.; Zhang, J.; Jiang, J.; Sheng, Y.; Wang, S.; Huang, A.; Tang, H. Up-regulated MicroRNA-181a induces carcinogenesis in hepatitis B virus-related hepatocellular carcinoma by targeting E2F5. *BMC Cancer* **2014**, *14*, 97, doi:10.1186/1471-2407-14-97.
53. Kim, R.; Emi, M.; Tanabe, K. Cancer immunoediting from immune surveillance to immune escape. *Immunology* **2007**, *121*, 1-14, doi:10.1111/j.1365-2567.2007.02587.x.
54. Hua, S.; Quan, Y.; Zhan, M.; Liao, H.; Li, Y.; Lu, L. miR-125b-5p inhibits cell proliferation, migration, and invasion in hepatocellular carcinoma via targeting TXNRD1. *Cancer Cell Int* **2019**, *19*, 203, doi:10.1186/s12935-019-0919-6.
55. Liu, W.; Hu, J.; Zhou, K.; Chen, F.; Wang, Z.; Liao, B.; Dai, Z.; Cao, Y.; Fan, J.; Zhou, J. Serum exosomal miR-125b is a novel prognostic marker for hepatocellular carcinoma. *Onco Targets Ther* **2017**, *10*, 3843-3851, doi:10.2147/OTT.S140062.
56. Kim, J.K.; Noh, J.H.; Jung, K.H.; Eun, J.W.; Bae, H.J.; Kim, M.G.; Chang, Y.G.; Shen, Q.; Park, W.S.; Lee, J.Y., et al. Sirtuin7 oncogenic potential in human hepatocellular carcinoma and its regulation by the tumor suppressors MiR-125a-5p and MiR-125b. *Hepatology* **2013**, *57*, 1055-1067, doi:10.1002/hep.26101.
57. Zhao, L.; Hu, K.; Cao, J.; Wang, P.; Li, J.; Zeng, K.; He, X.; Tu, P.F.; Tong, T.; Han, L. lncRNA miat functions as a ceRNA to upregulate sirt1 by sponging miR-22-3p in HCC cellular senescence. *Aging (Albany NY)* **2019**, *11*, 7098-7122, doi:10.18632/aging.102240.
58. Huang, Y.; Xiang, B.; Liu, Y.; Wang, Y.; Kan, H. LncRNA CDKN2B-AS1 promotes tumor growth and metastasis of human hepatocellular carcinoma by targeting let-7c-5p/NAP1L1 axis. *Cancer Lett* **2018**, *437*, 56-66, doi:10.1016/j.canlet.2018.08.024.

Figures

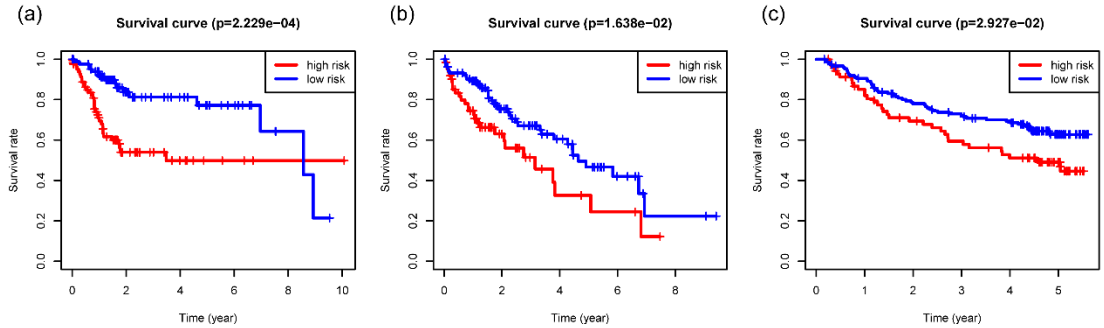

Figure 1. Performance of survival analysis demonstrating the association of the two-E2F signature score with the prognosis of HCC patients. (a, b, c) Kaplan–Meier survival curves for comparison of overall survival rates between patients in the low-risk group and the high-risk group for the TCGA training cohort (a), the TCGA validation cohort (b) and the external validation cohort from GEO database (c).

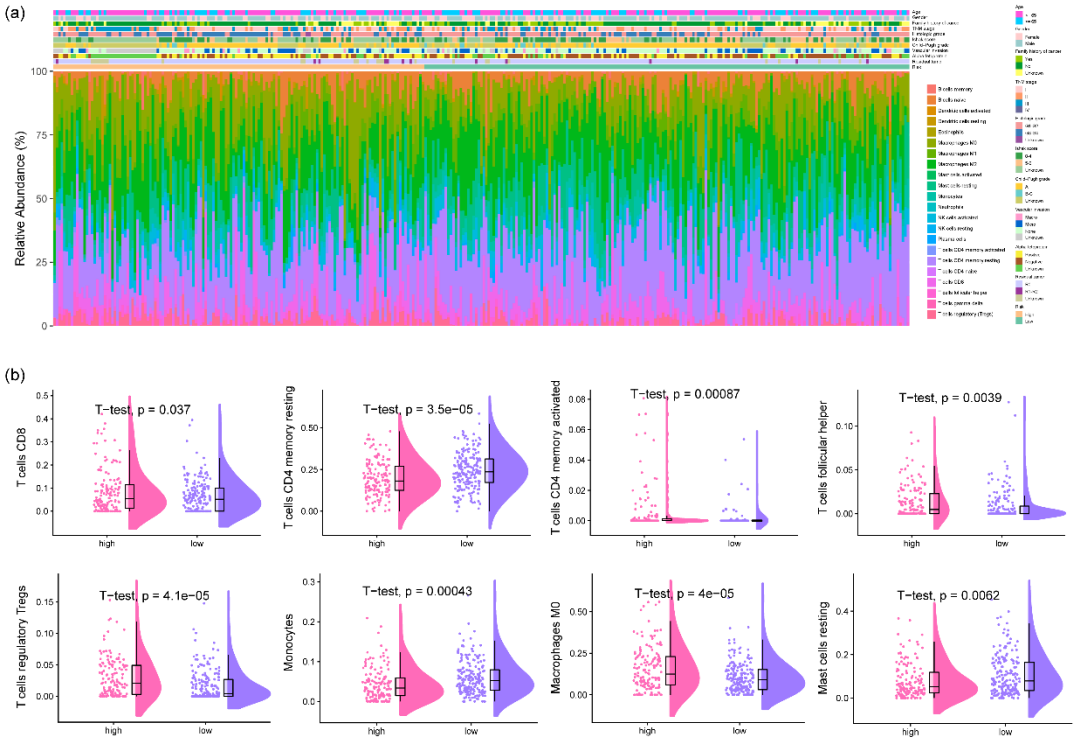

Figure 2. Full landscape of immune cell infiltration in HCC patients from the high-risk and low-risk groups. (a) Distribution in fractions of diverse tumor infiltrating immune cells in high-risk and low-risk HCC patients. (b) Comparisons of the proportions of different tumor infiltrating immune cells between the high-risk and low-risk patients. Only the comparisons with statistical significance were visualized.

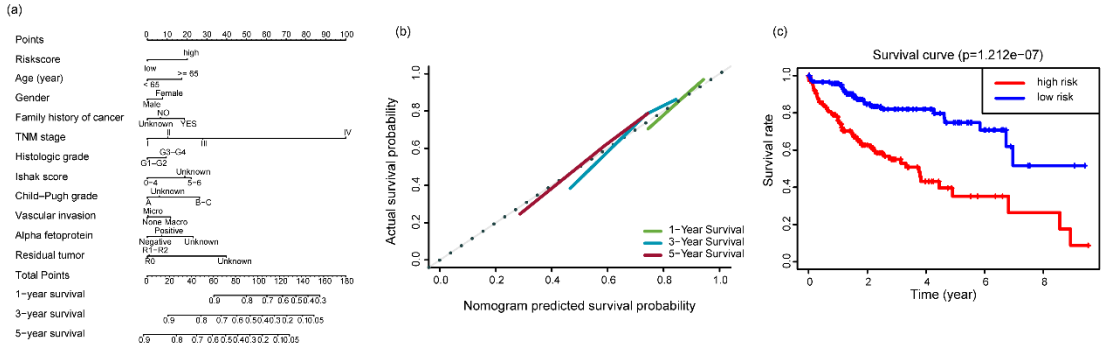

Figure 3. Construction and validation of the nomogram predicting the overall survival for HCC patients. (a) The nomogram predicting the 1-year, 3-year and 5-year overall survival probability of the HCC patients was created by integrating the two-E2F prognostic signature with the common clinical features. (b) The calibration plot showing the predicted and the actual observed survival rates for HCC patients at the 1-year, 3-year and 5-year time points. (c) Performance of the nomogram in predicting prognosis for the HCC patients.

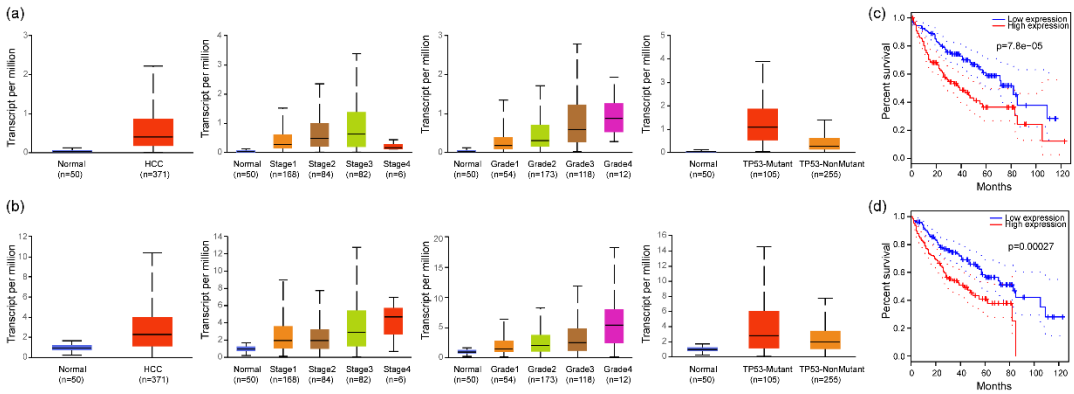

Figure 4. The expression levels and prognosis prediction of E2F2 and E2F5. (a, b) Association of the E2F2 (a) and E2F5 (b) expression levels with HCC and subclasses of HCC based on different clinicopathological parameters including stage, grade and TP53 status. (c, d) Kaplan-Meier plots for the survival analysis of E2F2 (c) and E2F5 (d), with patients being stratified according to the median expression value of the respective gene.

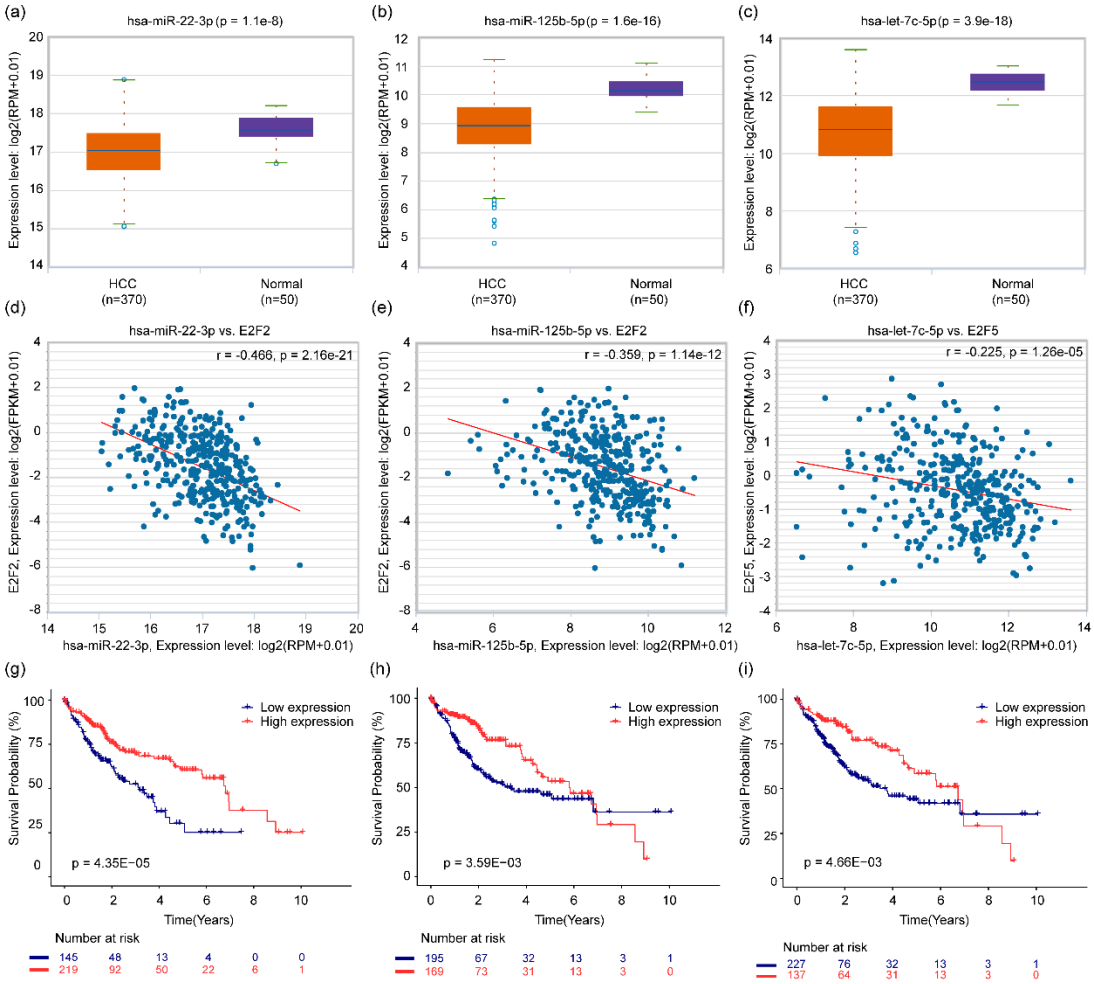

Figure 5. Characterization of the three key upstream miRNAs of E2F2 and E2F5. (a, b, c) Differential expression of hsa-miR-22-3p (a), hsa-miR-125b-5p (b) and hsa-let-7c-5p (c) between HCC patients and normal controls. (d, e, f) Negative correlations between the expression of hsa-miR-22-3p (d), hsa-miR-125b-5p (e) and hsa-let-7c-5p (f) and the respective target proteins. (g, h, i) Performance of the survival analysis displaying the capacity of hsa-miR-22-3p (g), hsa-miR-125b-5p (h) and hsa-let-7c-5p (i) in predicting prognosis of HCC patients.

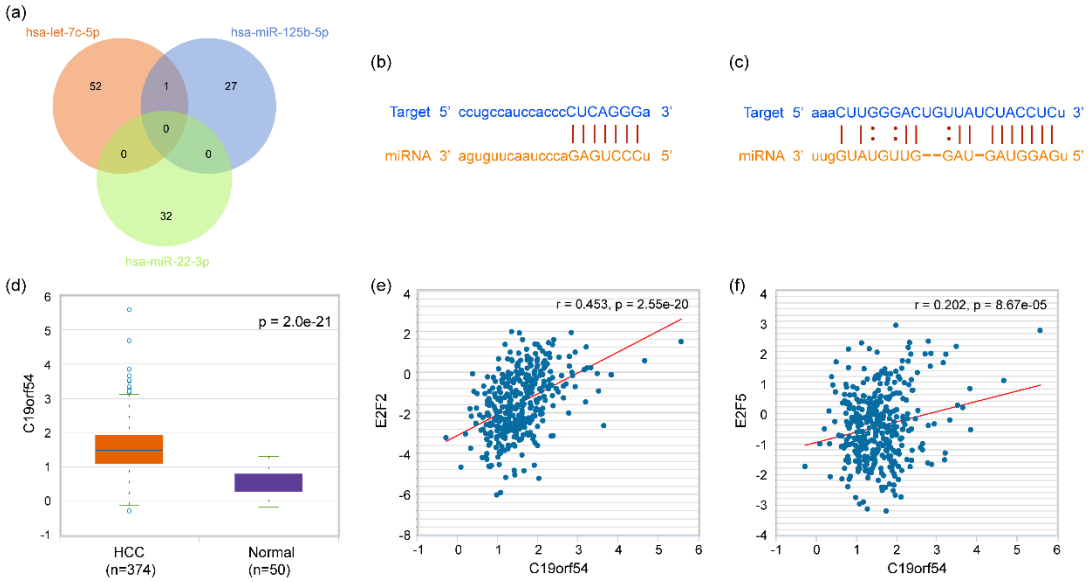

Figure 6. The common circular RNA circ-C19orf54 regulated by hsa-miR-125b-5p and hsa-let-7c-5p. (a) Venn diagrams displaying the overlap circRNA between the miRNAs. (b, c) The predicted binding sites of C19orf54 for hsa-miR-125b-5p (b) and hsa-let-7c-5p (c). (d) Differential expression of C19orf54 between HCC patients and normal controls. (e, f) The expression level of C19orf54 was positively correlated with the expression level of E2F2 (e) and E2F5 (f).

Tables

Table 1. Association between the risk score derived from the prognostic signature and the common clinicopathological factors.

		Risk score		
	Total (N=339)	**high (N=147)**	**low (N=192)**	**P-value**
Age (year)				
< 65	208 (61.4%)	97 (66.0%)	111 (57.8%)	0.156
≥ 65	131 (38.6%)	50 (34.0%)	81 (42.2%)	
Gender				
Male	231 (68.1%)	100 (68.0%)	131 (68.2%)	1
Female	108 (31.9%)	47 (32.0%)	61 (31.8%)	
Family history of cancer				
NO	196 (57.8%)	92 (62.6%)	104 (54.2%)	0.0697
YES	98 (28.9%)	33 (22.4%)	65 (33.9%)	
Unknown	45 (13.3%)	22 (15.0%)	23 (12.0%)	

TNM stage				
I	170 (50.1%)	63 (42.9%)	107 (55.7%)	0.0155
II	84 (24.8%)	40 (27.2%)	44 (22.9%)	
III	81 (23.9%)	44 (29.9%)	37 (19.3%)	
IV	4 (1.2%)	0 (0%)	4 (2.1%)	
Histologic grade				
G1–G2	212 (62.5%)	75 (51.0%)	137 (71.4%)	<0.001
G3–G4	125 (36.9%)	70 (47.6%)	55 (28.6%)	
Unknown	2 (0.6%)	2 (1.4%)	0 (0%)	
Ishak score				
0-4	124 (36.6%)	46 (31.3%)	78 (40.6%)	<0.001
5-6	74 (21.8%)	22 (15.0%)	52 (27.1%)	
Unknown	141 (41.6%)	79 (53.7%)	62 (32.3%)	
Child–Pugh grade				
A	207 (61.1%)	80 (54.4%)	127 (66.1%)	0.0193
B-C	21 (6.2%)	7 (4.8%)	14 (7.3%)	
Unknown	111 (32.7%)	60 (40.8%)	51 (26.6%)	
Vascular invasion				
None	193 (56.9%)	73 (49.7%)	120 (62.5%)	0.00389
Micro	84 (24.8%)	35 (23.8%)	49 (25.5%)	
Macro	14 (4.1%)	7 (4.8%)	7 (3.6%)	
Unknown	48 (14.2%)	32 (21.8%)	16 (8.3%)	
Alpha fetoprotein				
Negative	143 (42.2%)	42 (28.6%)	101 (52.6%)	<0.001
Positive	120 (35.4%)	66 (44.9%)	54 (28.1%)	
Unknown	76 (22.4%)	39 (26.5%)	37 (19.3%)	
Residual tumor				
R0	301 (88.8%)	123 (83.7%)	178 (92.7%)	0.0303
R1-R2	12 (3.5%)	7 (4.8%)	5 (2.6%)	
Unknown	26 (7.7%)	17 (11.6%)	9 (4.7%)	

Table 2. Univariate and multivariate analysis of overall survival.

	Univariate analysis		**Multivariate analysis**	
Variables	HR (95% CI)	***p***	HR (95% CI)	***p***
Age (≥ 65 vs. < 65)	1.23(0.85,1.78)	0.273	-	-
Gender (Female vs. Male)	1.26(0.87,1.84)	0.228	-	-
Family history of cancer (YES vs. NO)	1.14(0.76,1.69)	0.530	-	-
TNM stage (II vs. I)	1.42(0.87,2.32)	0.160	1.23(0.66,2.3)	0.507
TNM stage (III vs. I)	2.72(1.78,4.15)	**<0.001**	2.11(1.24,3.59)	**0.006**
TNM stage (IV vs. I)	5.44(1.68,17.63)	**0.005**	7.3(2.19,24.35)	**0.001**
Histologic grade (G3–G4 vs. G1–G2)	1.14(0.78,1.67)	0.489	-	-
Ishak score (5–6 vs. 0–4)	0.87(0.5,1.5)	0.612	-	-
Child–Pugh grade (B–C vs. A)	1.66(0.82,3.36)	0.159	-	-
Vascular invasion (Micro vs. none)	1.16(0.72,1.88)	0.539	0.97(0.56,1.69)	0.914
Vascular invasion (Macro vs. none)	2.52(1.14,5.58)	**0.023**	1.87(0.82,4.27)	0.135
Alpha fetoprotein (positive vs. negative)	1.45(0.92,2.28)	0.108	-	-
Residual tumor (R1–R2 vs. R0)	1.17(0.43,3.2)	0.754	-	-
Risk score (high vs. low)	2.17(1.5,3.15)	**<0.001**	1.86(1.19,2.9)	**0.007**