

## Research Paper

# Transcription factor E2F4 is an indicator of poor prognosis and is related to immune infiltration in hepatocellular carcinoma

Qiuxian Zheng<sup>1\*</sup>, Qiang Fu<sup>2\*</sup>, Jia Xu<sup>1\*</sup>, Xinyu Gu<sup>1</sup>, Haibo Zhou<sup>1</sup>, Chen Zhi<sup>1</sup>✉

1. State Key Laboratory for Diagnosis and Treatment of Infectious Diseases, National Clinical Research Center for Infectious Diseases, Collaborative Innovation Center for Diagnosis and Treatment of Infectious Diseases, The First Affiliated Hospital, College of Medicine, Zhejiang University, Hangzhou 310003, China.
2. School of Continuing Education, Zhejiang University, Hangzhou 310003, China.

\*These authors contributed equally to this work.

✉ Corresponding author: Prof. Chen Zhi, The First Affiliated Hospital, College of Medicine, Zhejiang University, No.79 Qingchun Road, Shangcheng District, Hangzhou 310003, Zhejiang Province, China; Tel: +86 0571-87236459; E-mail: zjuchenzhi@zju.edu.cn

© The author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0/>). See <http://ivyspring.com/terms> for full terms and conditions.

Received: 2020.08.06; Accepted: 2020.12.26; Published: 2021.01.21

## Abstract

**Background:** Recent studies have shown that the transcription factor E2F4 is involved in the progression of various tumors, but its expression and influence on immune cell infiltration and biological functions are largely unknown in hepatocellular carcinoma (HCC).

**Methods:** The Cancer Genome Atlas (TCGA) database, the Tumor Immune Estimation Resource (TIMER) and related online tools as well as a tissue microarray (TMA) were used for analyses in our study.

**Results:** E2F4 expression was elevated in HCC tumor tissue compared with adjacent normal tissue at both the mRNA and protein levels. Overexpression of E2F4 was markedly related to a poor prognosis in HCC patients. In addition, positively and negatively correlated significant genes of E2F4 were identified in HCC. Pathway enrichment analyses revealed that the top 100 positively correlated significant genes of E2F4 were closely related to nuclear splicing and degradation-related pathways. Furthermore, nine hub genes correlated with E2F4 expression were validated based on a protein-protein interaction (PPI) network. It was also demonstrated that E2F4 expression was negatively correlated to immune purity and positively correlated to immune cell infiltration.

**Conclusion:** E2F4 could serve as a novel biomarker for HCC diagnosis and prognosis prediction.

Key words: HCC, E2F4, hub genes, prognosis, immune infiltration

## Introduction

Hepatocellular carcinoma (HCC) has become the third-leading cause of cancer-related mortality worldwide [1, 2], with the devastating complication of chronic liver disease, due to the difficult detection of HCC during the early stages when the disease is curable [3]. Many patients with advanced HCC have rapid cell proliferation, metastasis, and insensitivity to chemotherapy with an abysmal prognosis. Consequently, the majority of HCC patients are diagnosed after the tumor has grown to a considerable size [4]. Therefore, novel biomarkers

provide an approach for early detection, which is urgently needed to improve the prognosis of HCC patients.

E2F, a group of genes encoding a transcription factor protein family, act as regulators of cell differentiation, response to DNA damage, and cell life cycle [5, 6]. It has been shown that compared to normal tissues, E2F family genes are highly expressed in several types of tumor tissue, including HCC [7]. The E2F family is one of the critical tumor-suppressor/oncogene regulators are also involved in

regulating telomerase expression and drug sensitivity [8, 9]. Hypo-methylation, amplification, and tumor protein p53 (TP53) mutations have a close relationship with the E2F family genes in many cancer types including HCC[6, 10]. As a well-characterized potent transcriptional inhibitor and E2F family member, E2F4 is a crucial factor that promotes tumor growth via recruiting pocket proteins, binding with retinoblastoma (RB) family proteins and functioning with histone-modifying enzymes[11], which is crucial for the regulation of differentiation and cell cycle arrest[12]. Certain evidence has illustrated that E2F4 is highly enriched in leucine/isoleucine-rich hydrophobic nuclear export signals, which cannot activate E2F-responsive genes *in vivo* [13]. Researchers have shown that E2F4 is involved in the tumorigenesis of Burkitt lymphoma [14], cervical cancer [15], colorectal cancer [16], and acute myeloid leukemia [17]. However, the expression and specific roles of E2F4, including the protein regulation network, remain unclear in HCC.

A growing body of evidence also shows that E2F4 is remarkably correlated with tumor immune cell infiltration [18]. E2F4 functions as an inducible protein in the immune response of eukaryotes. In mammals, E2F4 is a critical member of transcriptional factors involved in innate immune responses, such as expression of toll-like receptor 8 (TLR8) and CD14, and the downstream signal transducer and activator of transcription (STAT1) pathway [18]. However, as an immuno-suppressive protein, the regulatory function of E2F4 in HCC remains unknown.

To enhance our understanding of E2F4 crosstalk with immune cells, we determined the relationship between E2F4 and programmed death-ligand 1 (PD-L1) expression, and also validated the relationship between E2F4 gene expression and immune cell infiltration. We analyzed the HCC cohort from the TCGA dataset using multiple online resources. We also investigated the expression level, clinical characteristics, mutations, and differentially expressed genes (DEGs) of E2F4. With these data, we assessed the signaling pathways of DEGs and contrasted the protein-protein interaction (PPI) network to identify hub genes. Finally, the expression of E2F4 in HCC and its association with immune cell infiltration were also explored.

## Materials and Methods

### Data collection and processing

The RNA sequencing (RNA-seq) and corresponding clinical data from pan-cancer patients as well as HCC patients were downloaded from The Cancer Genome Atlas Liver Hepatocellular Carcinoma (TCGA-LIHC) data collection (<https://xena.ucsf.edu/>) as previously described

[19]. Three hundred seventy-four liver cancer tissues and 50 adjacent tissues were included in the HCC patient cohort. E2F4 gene expression profiles were further analyzed by characteristics according to age, race, tumor grade, tumor stages, TP53 gene mutation status, and lymph node metastasis status. E2F4 gene expression was determined in pan-cancers and HCC patients. The pan-cancer cohort gene profiles were also downloaded from the publicly available TCGA database, including tumor and matched non-tumor tissues. The E2F4 gene expression level among various diseases was then estimated. The expression of E2F4 across the pan-cancer data was assessed via the UALCAN website (<http://ualcan.path.uab.edu/>) [20].

### Overall clinical survival of E2F4 gene-related HCC patients

The survival curves of HCC cohort patients were analyzed online via The Kaplan Meier plotter (<https://kmplot.com/>) based on TCGA analysis as described previously [21, 22]. The overall survival (OS), recurrence-free survival (RFS), progression-free survival (PFS), and disease-specific survival (DSS) were the designated endpoints. Patients were divided into low and high expression groups based on the median expression of E2F4. Comparisons were made between the two groups for each type of survival curve.

### Mutation types of E2F4 gene

The gene mutations and co-expression of E2F4 were computed and analyzed using CBio-Cancer Genomics Portal (cBioPortal) databases (<https://www.cbioportal.org/>) [23]. Searching the term “E2F4” enabled the acquisition of the full mutation distribution across all tumor and non-tumor tissues. After selecting “HCC”, the mutation frequencies and mutation types that occur in HCC samples were obtained.

### E2F related immune infiltration

Tumor Immune Estimation Resource (TIMER, <https://cistrome.shinyapps.io/timer>) is a database resource for systematic evaluations that contains information regarding the clinical impact of immune infiltration among various types of cancer [24]. Immune cells, such as B cells, T cells, neutrophils, and antigen-presenting cells (APCs) were included. The microenvironment in tumor tissues was estimated via a novel statistical algorithm. Furthermore, considering that TP53 is the most frequently mutated gene in HCC, the correlation between TP53 mutation status and immune cell infiltration was analyzed via the TIMER website.

## Identification of differentially expressed genes

LinkedOmics (<http://www.linkedomics.org/>) contains various clinical data for pan-cancer types, and all patients were included as described in a previous study [25]. To determine candidate biomarkers, gene expression profiles were analyzed. To further identify DEGs, associations between billions of attribute pairs for each cancer cohort were analyzed and visualized. The top 10 positively and negatively co-related genes were selected.

## E2F4-related hub gene selection and validation

The PPI network and module analysis were performed for further exploration. The top nine genes identified by degree were considered hub genes and were noted for further study. Additionally, the Gene Expression Profiling Interactive Analysis [26] (GEPIA, <http://gepia.cancer-pku.cn/>) and Cytoscape software (<https://cytoscape.org/>) were used to determine hub genes and investigate the correlation between hub genes and E2F4 expression.

## Pathway network construction

To decipher the potential molecular mechanisms and explore the functional annotation of E2F4-related DEGs in HCC genomic profiles, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were performed.

## Clinical tissues

Tissue microarray (TMA) containing 90 HCC and 90 normal liver tissues was purchased from Outdo Biotech (Shanghai, China). All of the clinical information including age, gender, grade, TNM stage, as well as survival time, was collected and the details are presented in Table 1.

## Immunohistochemistry (IHC) analysis

The IHC staining pattern of E2F4 was evaluated. Briefly, after incubation with the E2F4 primary antibody (dilution 1:300, [EPR8259] (ab150360) Abcam, Cambridge, MA), the tissue microarray was probed with biotinylated goat anti-rabbit secondary antibody rabbit IgG, (Dilution: 1:10,000, Cell Signaling Technologies, Inc.), and the remainder of the process was carried out according to a previously published study [27].

## Statistical analysis

A *t*-test with a *p* value < 0.05 was applied to identify significance with statistical differences. The OS analysis was performed using R programming of "survival" and "survive" packages online. The OS, RFS, PFS, and DSS were separated and calculated by the median gene expression level of E2F4. Spearman's correlation was applied to identify the relationship

between E2F4 and associated genes and pathways. A *p* value < 0.05 was used to identify significant differences.

## Results

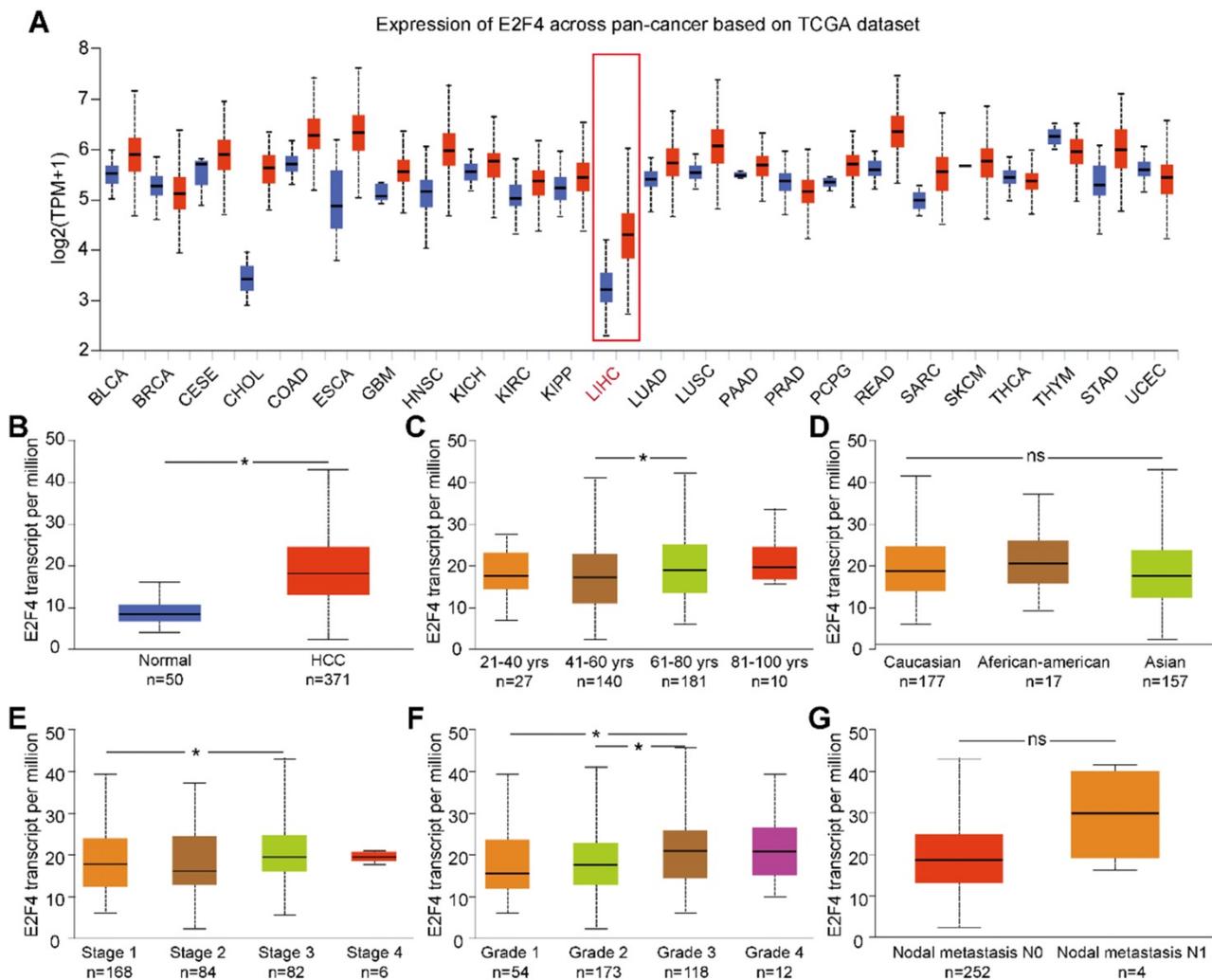
### E2F4 was highly expressed in pan-cancers, and the E2F4 gene was differently expressed according to the clinical characteristics of HCC

UALCAN was used to analyze the expression of E2F4 mRNA from the pan-cancer data. The results showed that the transcription of E2F4 was increased in most cancers (Figure 1A). Additionally, the transcriptional level of E2F4 in the TCGA-LIHC cohort was significantly higher compared with normal tissue (Figure 1B). To further investigate the levels of E2F4 mRNA in HCC patients, the clinicopathological characteristics of HCC patients were analyzed. In this study, a slight but statistically significant difference was observed between stages 1 and 3. In regards to the tumor grade, significant differences in E2F4 transcription were found between both grade 1 and grade 2 when compared to grade 3 (all *p* < 0.05; Figure 1E and 1F). However, the E2F4 mRNA showed no significant differences in regard to age, race, or node metastasis status among patients with HCC (Figure 1C, 1D and 1G). These findings suggest that E2F4 expression was upregulated in HCC tissues, and it has potential to be a novel biomarker for HCC diagnosis.

**Table 1.** Correlation between E2F4 expression and clinicopathological characteristics of patients with HCC

Variables	E2F4 expression		Total	$\chi^2$	p value
	Low	High			
Age (years)				0.262	0.608
	≤ 50	15	25	40	
	> 50	21	28	49	
Sex				1.116	0.291
	Female	2	8	10	
	Male	34	45	79	
Grade				0.597	0.44
	1/2	22	28	50	
	3	14	25	39	
TNM stage				0.187	0.665
	I	26	36	62	
	II/III	10	17	27	
Cirrhosis				0	1
	Negative	4	5	9	
	Positive	32	47	79	
AFP				1.154	0.283
	Negative	18	20	38	
	Positive	18	32	50	
PD-L1				4.751	0.029
	Negative	13	11	24	
	Positive	17	42	59	

Abbreviations: HCC, hepatocellular carcinoma; TNM, tumor node metastasis; AFP, alpha-fetoprotein; PD-L1, programmed death-ligand 1.



**Figure 1. E2F4 gene expression and the clinical characteristics associated with E2F4 expression and clinical stage based on data from TCGA database.** (a) Comparison of E2F4 expression between different cancer types and normal tissues in the TCGA dataset. (b) E2F4 expression in HCC tumor tissues compared to non-tumor tissues. (c) E2F4 expression in HCC tumor tissues among different age groups. (d) E2F4 expression in HCC based on different races. (e) E2F4 expression based on the tumor stage in HCC. (f) Comparison between different HCC tumor grades and E2F4 gene expression. (g) E2F4 expression with nodal metastasis N0 and N1. Chi-square test analysis was applied in this study and we recognised p value <0.05 as statistically significant. **Abbreviations:** LIHC, liver hepatocellular carcinoma; HCC: hepatocellular carcinoma; TCGA: The Cancer Genome Atlas; TPM: transcripts per million.

## Elevated E2F4 expression impacts the clinical prognosis of patients with HCC

To further explore the relationship between E2F4 expression and clinical outcomes in patients with HCC, a Kaplan-Meier plot was used to evaluate patient prognosis in relation to different E2F4 expression levels. The OS of HCC patients with high E2F4 expression was significantly worse than those with low E2F4 expression (hazard ratio [HR] = 1.5 [95% confidence interval (CI): 1.03-2.17], log rank  $p = 0.032$ ; **Figure 2A**). Moreover, patients with high expression of E2F4 had worse DSS (HR = 1.71 [95% CI: 1.05-2.78], log rank  $p = 0.029$ ; **Figure 2B**). There were no significant differences in PFS (HR = 1.3 [95% CI: 0.97-1.76], log rank  $p = 0.082$ ; **Figure 2C**) or RFS (HR = 1.39 [95% CI: 1-1.95], log rank  $p = 0.052$ ; **Figure 2D**). Collectively, elevated E2F4 expression was

significantly related to a worse prognosis for patients, indicating the E2F4 may be able to serve as a prognostic factor.

## E2F4 expression was increased in HCC tissues and related to poorer prognosis of patients

The association between E2F4 protein expression and clinical characteristics of patients with HCC were assessed. Representative micrographs reveal E2F4 IHC staining patterns with different staining scores in HCC and paired normal liver tissues, and the staining patterns were categorized into three groups, 1+, 2+ and 3+, according to staining intensity (**Figure 3A and 3B**). As shown in the **Figure 3C**, E2F4 expression was increased in HCC tissues when compared to the paired liver tissues ( $p = 0.036$ , **Figure 3C**), which was consistent with the results from the TCGA cohort (**Figure 1B**). In addition, we found that elevated E2F4

expression was positively correlated with the PD-L1 positive subgroup ( $p = 0.029$ , **Figure 3D**). There was no significant correlation among E2F4 expression and age, gender, grade, TNM stage, cirrhosis, or alpha-fetoprotein (AFP) levels (**Table 1**). Moreover, we assessed the correlation of E2F4 expression with survival of patients with HCC. Kaplan-Meier curves revealed that patients with increased E2F4 expression had shorter OS times compared with those who had reduced E2F4 expression (**Figure 3E**). Univariable and multivariable Cox regression analyses were adopted to determine the risk factors for OS prognosis. Univariable analysis demonstrated that advanced grade, late TNM stage, and high E2F4 expression were potential risk factors for decreased OS in HCC. Through further analysis, high E2F4 expression was found to be an independent predictor of OS according to multivariable analysis (**Table 2**). Taken together, these findings demonstrated that up-regulated E2F4 expression could service as a novel predictor for prognosis in HCC.

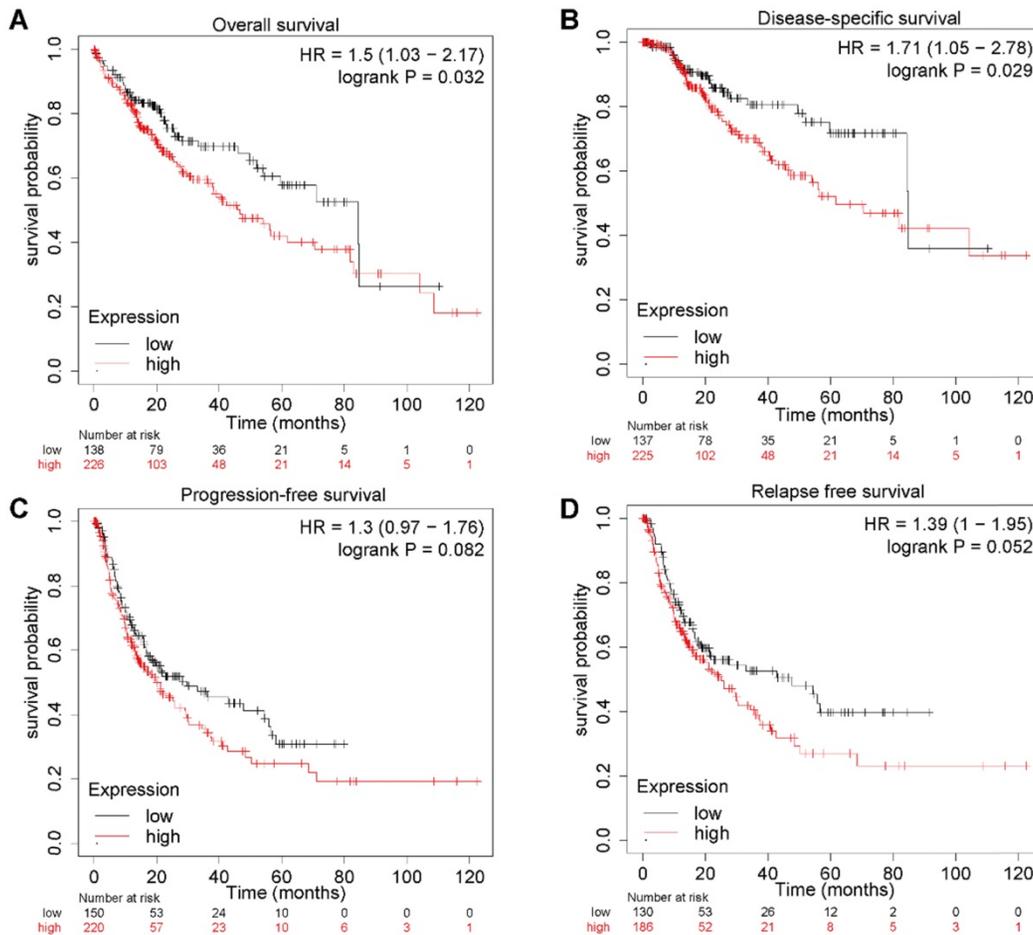
**Table 2.** Univariate and multivariate analyses of the factors correlated with overall survival of patients with HCC

	Univariate analysis			Multivariate analysis		
	HR	95% CI	p value	HR	95% CI	p value
Age	1.383	0.753-2.54	0.296			
Sex	1.438	0.445-4.652	0.544			
Grade	1.814	0.999-3.293	0.05	1.74	0.948-3.194	0.074
TNM stage	1.783	1.045-3.043	0.034	1.655	0.949-2.885	0.076
<b>E2F4 expression</b>	<b>2.071</b>	<b>1.096-3.917</b>	<b>0.025</b>	<b>2.121</b>	<b>1.119-4.023</b>	<b>0.021</b>

Abbreviations: HCC: hepatocellular carcinoma; HR, hazard ratio; CI, confidence interval; TNM, tumor node metastasis.

**The mutation characteristics of E2F4 and the correlation of mutated genes**

The mutation characteristics of E2F4 were further investigated, and an analysis using the cBioPortal was performed [28]. Results indicated that the frequency of somatic mutations in E2F4 was 0.3% in HCC (**Figure 4A**). Moreover, the correlation between E2F4 expression with the mutations and co-expressed genes was explored. We obtained eight highly co-expressed genes in HCC based on the TCGA database (**Table 3**). As shown in **Figure 4B-4I**,



**Figure 2. E2F4 gene expression and clinical outcomes.** (a) OS curves showing stratification of HCC patients in the TCGA dataset based on E2F4 expression levels. (b) DSS curves of patients with either high or low E2F4 gene expression. (c) PFS curves of patients with either high or low E2F4 gene expression. (d) RFS curves of patients with either high or low E2F4 gene expression. **Abbreviations:** OS, overall survival; TCGA, The Cancer Genome Atlas; DSS, disease-specific survival; PFS, progression-free survival; RFS, relapse-free survival.

E2F4 expression was significantly positively associated with the following genes: enhancer of mRNA decapping 4 (*EDC4*; Pearson correlation coefficient [PCC] 0.69,  $p = 1.41e-49$ ), IST1 factor associated with ESCRT-III (*IST1*; PCC 0.64,  $p = 1.03e-41$ ), nucleoporin 93 (*NUP93*; PCC 0.73,  $p = 6.94e-60$ ), U6 snRNA biogenesis phosphodiesterase 1 (*USB*; PCC 0.54,  $p = 2.80e-27$ ), solute carrier family 7 member 6 (*SLC7A6*; PCC 0.58,  $p = 7.03e-33$ ), nuclear transport factor 2 (*NUTF2*; PCC 0.57,  $p = 1.32e-31$ ), NEDD8 activating enzyme E1 subunit 1 (*NAE1*; PCC 0.58,  $p = 1.76e-34$ ), and adrenocortical dysplasia protein homolog (*ACD*; PCC 0.51,  $p = 1.17e-24$ ). These results provided the mutation characteristics of E2F4 and the distribution of the genes co-expressed with E2F4 in HCC, implying that E2F4 gene expression has close co-relationship with other genes expression.

**Table 3.** Top eight genes co-expressed with E2F4.

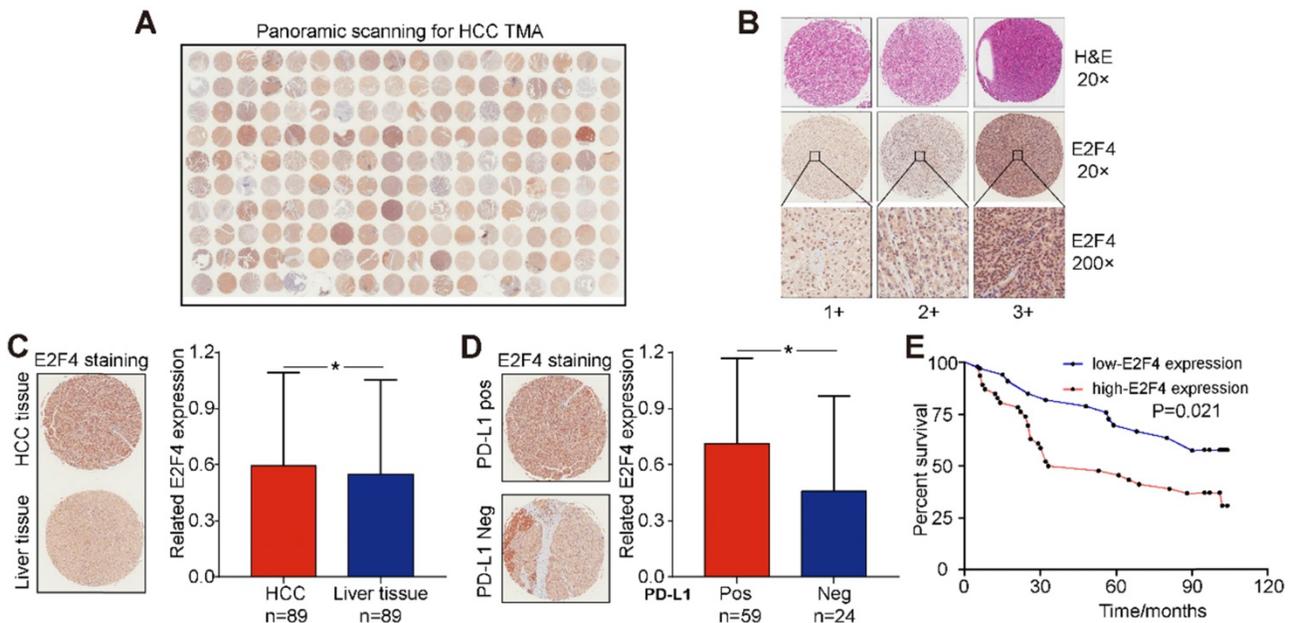
Correlated Gene	Cytoband	Spearman's Correlation	p value
EDC4	16q22.1	0.697	6.90E-52
IST1	16q22.2	0.655	5.47E-44
NUP93	16q13	0.654	8.89E-44
SLC7A6	16q22.1	0.634	1.44E-40
NUTF2	16q22.1	0.576	4.04E-32
NAE1	16q22.1	0.568	3.74E-31
ACD	16q22.1	0.567	5.84E-31
USB1	16q21	0.566	7.31E-31

Abbreviations: EDC4, enhancer of mRNA decapping 4; IST1, IST1 factor associated with ESCRT-III; NUP93, nucleoporin 93; SLC7A6, solute carrier family 7 member 6; NUTF2, nuclear transport factor 2; NAE1, NEDD8 activating enzyme E1 subunit 1; ACD, adrenocortical dysplasia protein homolog; USB1, U6 snRNA biogenesis phosphodiesterase 1.

**DEGs of E2F4 in HCC and selection of hub genes**

To further validate the biological activities of E2F4 in HCC, the E2F4 DEGs were evaluated in HCC. The volcano map identified E2F4-related DEGs with positively related genes on the right of the plot and negatively related genes on the left of the plot (**Figure 5A**). Additionally, the heatmaps of the top 10 negatively related genes identified were *CYP4A11*, *G6PC*, *ALDH5A1*, *TMEM56*, *INSIG1*, *ACOX1*, *SEPSECS*, *ABCA6*, *SLC10A1*, and *HMGCS2* (**Figure 5B**). The top 10 positively related genes were *EDC4*, *KIAA0714*, *NUP93*, *SLC7A6*, *CIZ1*, *PRMT1*, *NIP7*, *NUF2*, *NAE1*, and *DHX38* (**Figure 5C**).

To determine the relationship of the top 100 positively related genes of E2F4 in HCC, a PPI network was established. As shown in **Figure 5D**, frequent interaction among the top 100 genes have close relationships with E2F4 expression. After calculating by degree using Cytoscape software, we obtained nine hub genes that revealed the closest relationships. The nine hub genes were *HDX38*, *NUP37*, *NUP93*, *PPP2P1A*, *RNPS1*, *RUVBL1*, *SF3B3*, *U2AF2*, and *UBE2IR* (**Figure 5E**). The expression of E2F4 was significantly associated with the expression of the hub genes (**Supplementary Figure 1**), especially for *SF3B3* ( $R = 0.81$ ). Based on these results, the relationship between E2F4 and hub genes was a novel reason for tumor progression. Our results demonstrated that E2F4 expression was correlated



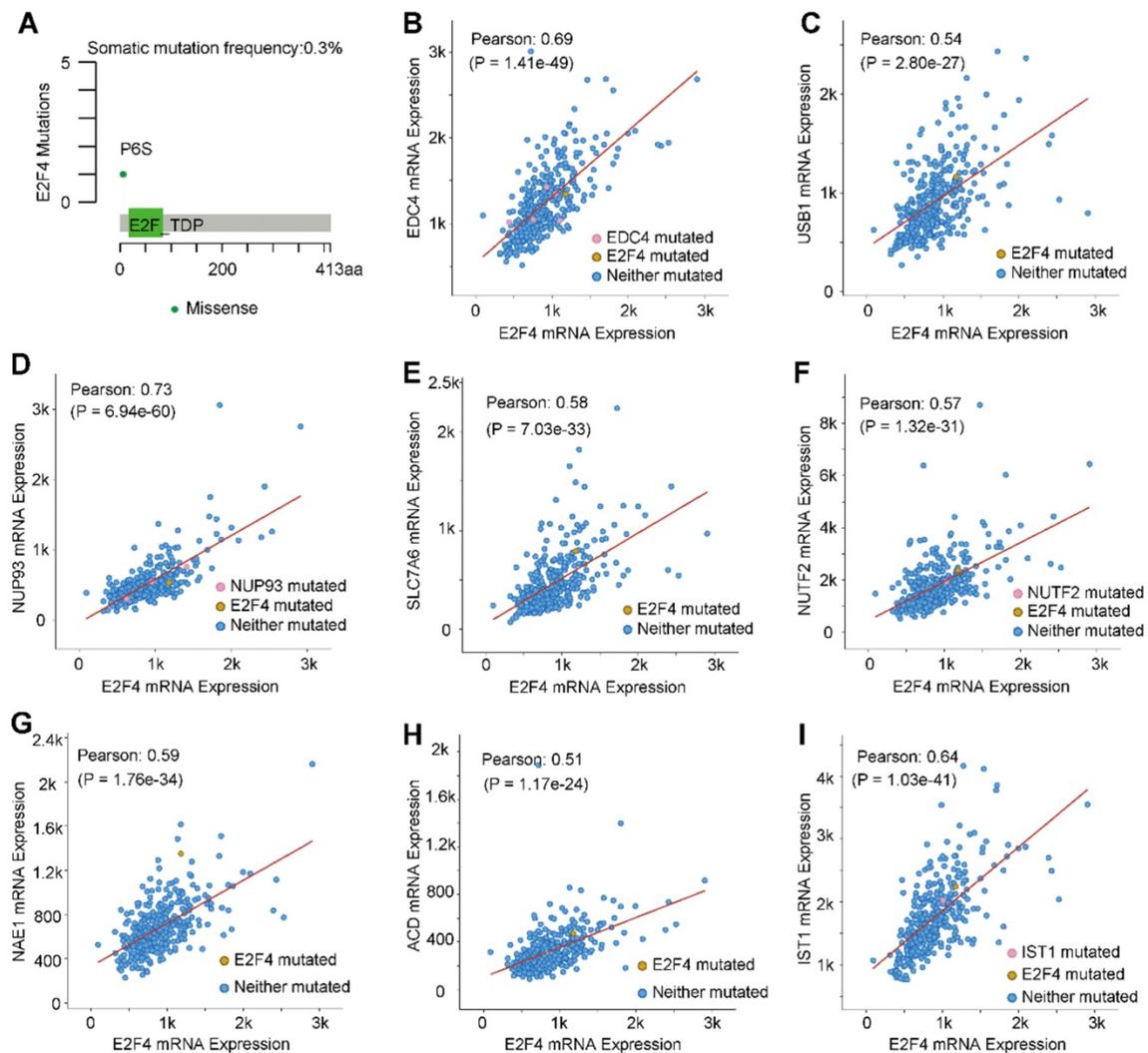
**Figure 3.** E2F4 expression was elevated in HCC tissues and high E2F4 expression was associated with poorer OS. (a) Representative E2F4 staining in HCC and normal tissues. (b) Representative E2F4 and immunohistochemical staining patterns with different staining scores in HCC tissues. (c) Chi-square test analysis was applied, and p value <0.05 was identified as statistically significant. Expression of E2F4 protein in HCC tissue compared to normal liver tissue. ( $p = 0.036$ ). (d) Chi-square test analysis. P value <0.05 is statistically significant. Association between up-regulated E2F4 expression and positive PD-L1 staining. ( $p = 0.029$ ,  $\chi^2=4.751$ ). (e) E2F4 expression associated with prognosis. **Abbreviations:** OS, overall survival; HCC, hepatocellular carcinoma; PD-L1, programmed death-ligand 1.

with numerous crucial genes involved in many biological regulatory activities such as the cell cycle, protein modification, cellular transcription, and glycolipid metabolism.

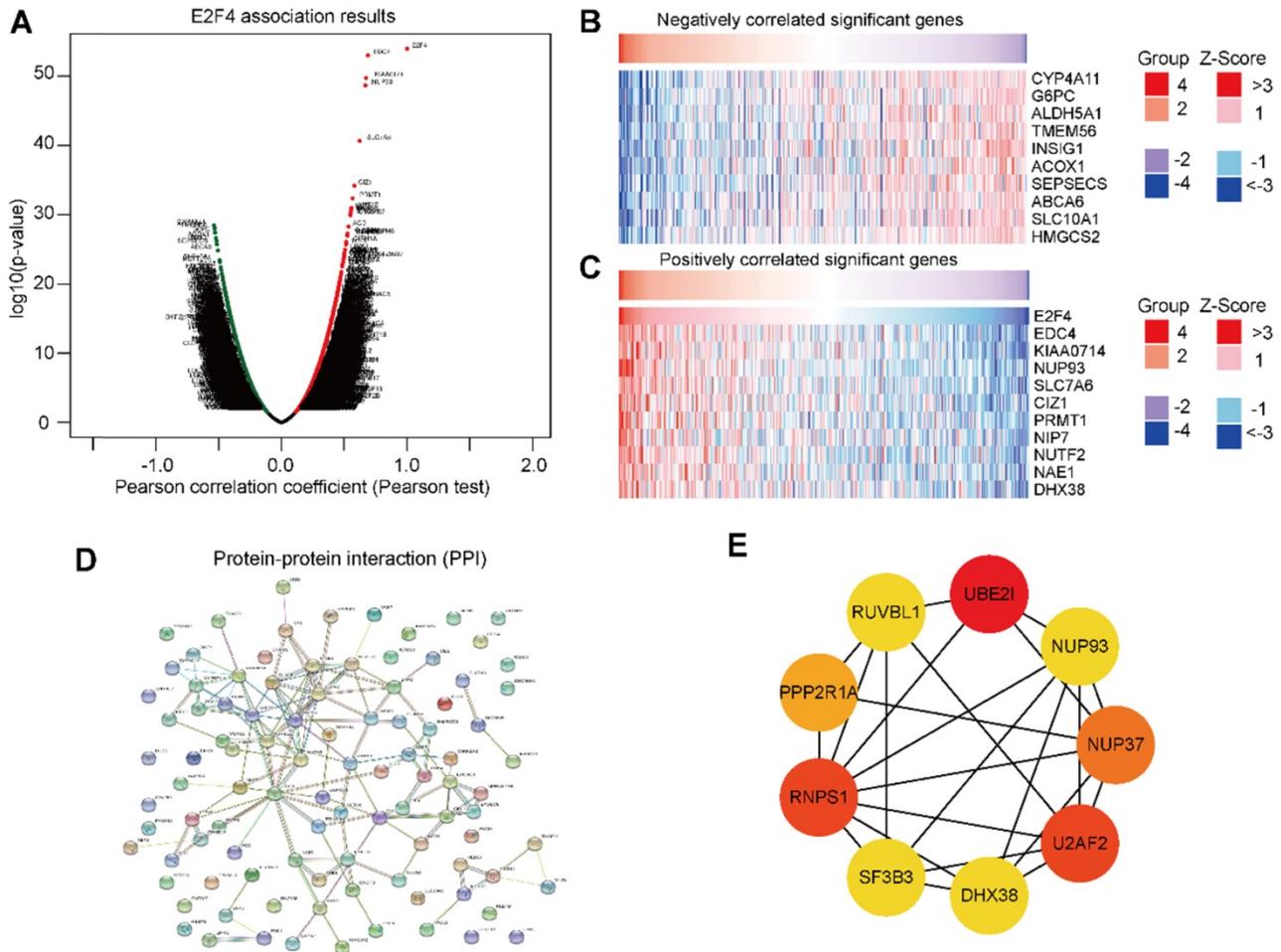
### E2F4 related pathways

To investigate the regulation of signaling pathways associated with E2F4, the top 100 significantly related genes were used in KEGG and GO pathway enrichment analyses. GO results revealed six significant biological processes (BP), nine cellular components (CC), and three molecular functions (MF; **Figure 6A**). More specifically, E2F4-related pathways involved in BPs included mRNA splicing via spliceosome, regulation of mitophagy, and positive regulation of protein targeting the mitochondrion. Negative regulation of mRNA splicing via spliceosome revealed that E2F4 actively participates in nucleic acid, mitochondria,

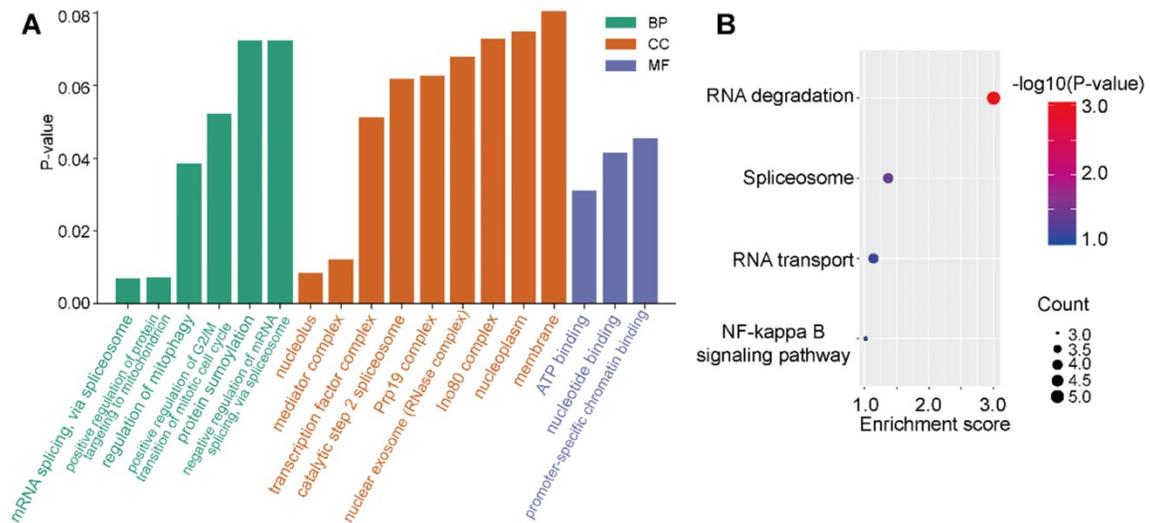
and accounting-related protein regulation. In CC analysis, the nucleolus, mediator complex associated pathways were significantly correlated with E2F4 function. In the MF subgroup, nucleotide binding, promoter-specific chromatin binding, and ATP binding pathways indicated significantly positive affiliation to E2F4 gene regulation (**Figure 6A**). Moreover, according to KEGG analysis, RNA degradation suggested that the E2F4 gene plays a crucial role in cellular biology and the E2F4 gene was positively associated with RNA degradation and negatively associated to RNA transport and the spliceosome (**Figure 6B**). These results collectively demonstrated that E2F4 has a wide range of effects on the production of the global transcriptome and the regulation of genes and pathways involved in mitochondrial metabolism.



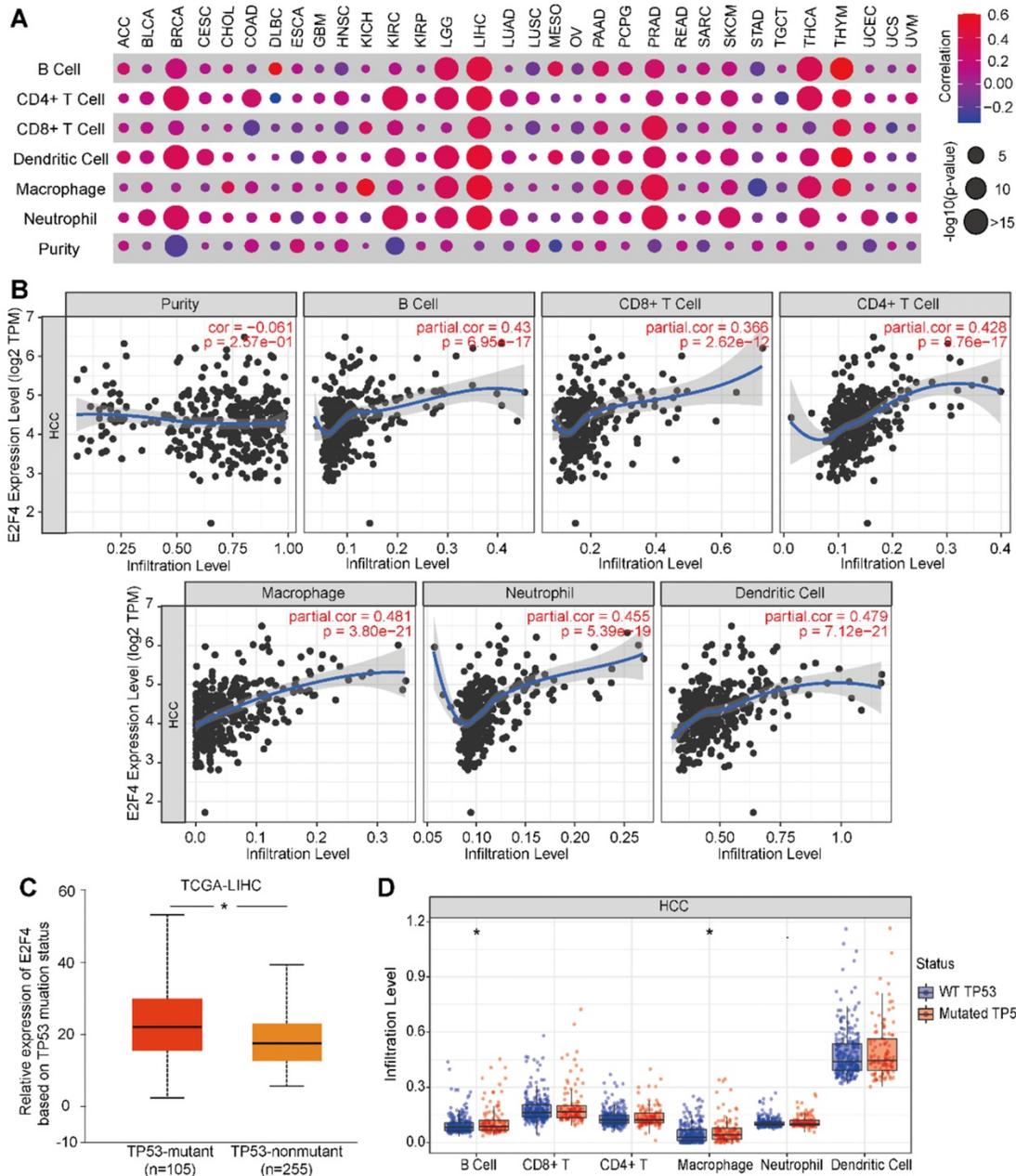
**Figure 4. E2F4 gene mutations and genes that are co-expressed in HCC. (a)** E2F4 gene somatic mutation frequency. **(b-i)** The relationship between a mutation in E2F4 and a mutation also occurring in (b) EDC4, (c) USB1, (d) NUP93, (e) SLC7A6, (f) NUTF2, (g) NAE1, (h) ACD, and (i) IST1. **Abbreviations:** HCC, hepatocellular carcinoma; Pearson, Pearson's correlation coefficient; EDC4, enhancer of mRNA decapping 4; USB1, U6 snRNA biogenesis phosphodiesterase 1; NUP93, nucleoporin 93; SLC7A6, solute carrier family 7 member 6; NUTF2, nuclear transport factor 2; NAE1, NEDD8 activating enzyme E1 subunit 1; ACD, adrenocortical dysplasia protein homolog; IST1, IST1 factor associated with ESCRT-III.



**Figure 5. Genes associated with E2F4 expression in HCC.** (a) Volcano plot showing the positively and negatively DEGs in TCGA. (b-c) The top 10 negatively and positively correlated genes with E2F4. (d) The PPI network of E2F4 expression. (e) The nine hub genes of E2F4. **Abbreviations:** HCC, hepatocellular carcinoma, DEGs, differentially expressed genes; TCGA, The Cancer Genome Atlas; PPI, protein-protein interaction. CYP4A11, specificity of cytochrome P450 4A11; G6PC, glucose-6-phosphatase (G6Pase) catalytic; ALDH5A1, Aldehyde dehydrogenase 5 family, member A1; INSIG1, insulin-induced gene 1; ACOX1, acyl Coenzyme A oxidase 1; SEPSECS, spontaneous excitatory postsynaptic currents; ABCA6, ATP-binding cassette subfamily A member 6; SLC10A1 (NTCP), Na-taurocholate cotransporting polypeptide; HMGCS2, 3-hydroxy-3-methylglutaryl-CoA synthase 2; EDC4, Enhancer of mRNA decapping protein 4; NUP93, Nucleoporin 93; SLC7A6, CIZ1, Cip1-interacting zinc finger protein 1; PRMT1, protein arginine N-methyltransferase 1; NIP7, nodulin 26-like intrinsic protein7; NUTF2, nuclear transport factor 2; NAE1, NEDD8-activating enzyme 1; DHX38, DEAH-Box Helicase 38. RUVBL1, RuvB-like protein 1; UBE2L, ubiquitin-conjugating enzyme E2L, PPP2R1A, Protein Phosphatase 2 Scaffold Subunit Aalpha; RNPS1, RNA-binding protein with serine-rich domain 1; SF3B3, U2AF2, U2 small nuclear RNA auxiliary factor 2, NUP37, Nucleoporin 37.



**Figure 6. Distribution of E2F4 associated pathways.** (a) Significantly enriched GO terms associated with BP, CC, and MF annotations in the TCGA-LIHC cohort. (b) Enrichment analysis of the E2F4 related pathway, where red indicates a significantly positively correlated pathways. **Abbreviations:** GO, gene ontology; BP, biological process; CC, cellular component; MF, molecular function; TCGA-LIHC, The Cancer Genome Atlas – liver hepatocellular carcinoma.



**Figure 7. The relevance of E2F4 gene expression and immune infiltration of the HCC cohort in the TIMER database. (a)** The correlation of E2F4 expression and immune cell infiltration across 32 solid cancers. **(b)** Correlation of immune purity with E2F4 gene expression. Immune infiltration levels in HCC of B cells, CD4+ T cells, CD8+ T cells, neutrophils, macrophages, and dendritic cells. **(c)** E2F4 expression correlated with TP53 mutation. **(d)** The affect of TP53 mutation on the infiltration of the immune cells. **Abbreviations:** HCC, hepatocellular carcinoma; TIMER, Tumor Immune Estimation Resource; TP53, tumor protein p53.

### E2F4 affects tumor purity, immune infiltration, and the immune signature of HCC

Next, we evaluated if E2F4 expression was associated with tumor immune status across 32 solid cancer types. E2F4 expression was found to be significantly associated with immune cell infiltration among several cancers, especially in HCC (Figure 7A). In HCC, E2F4 was positively correlated with immune cell infiltration by cells including B cells (partial. cor = 0.43, p = 6.95e-17), CD4+ T cells (partial. cor = 0.428, p = 9.76e-17), CD8+ T cells (partial. cor =

0.366, p = 2.62e-12), macrophages (partial. cor = 0.481, p = 3.80e-21), neutrophils (partial. cor = 0.455, p = 5.39e-19), and dendritic cells (partial. cor = 0.479, p = 7.12e-21). However, E2F4 expression was negatively related to immune purity (Figure 7B). Considering the high frequent of TP53 mutations in HCC, we investigated the effect of mutated TP53 on E2F4 expression. As shown in Figure 7C, we observed that E2F4 expression was elevated in the TP53-mutant group (n = 105) compared with the TP53 non-mutant group (n = 255). The TP53 mutation status was further evaluated regarding the level of immune cell

infiltration. In this study, it was discovered that B cells and macrophages were increased with TP53 mutation compared to wild-type TP53 (**Figure 7D**). Our results demonstrated that TP53 mutation might play an essential regulatory role in tumor immune cell infiltration.

## Discussion

HCC is a worldwide threat with significant morbidity and mortality based on the stage of the disease at presentation [29]. Current treatment strategies such as resection, chemotherapy, radiation, immunotherapy, and gene therapy are improving, but recurrence of illness due to therapy resistance results in high mortality. And the five-year recurrence rate after resection is up to 60% to 70% [30]. Most HCC patients suffer from a clinically poor prognosis [31]. Therefore, it is urgent to find out effective prognostic biomarkers and comprehensively acknowledged the genomic associated characteristics to predict the prognosis of patients with HCC before resection. Hence, the selection of optimum candidate genes for early detection and targeted therapy is especially important. The origin of HCC is a multistep process with a highly progressive accumulation of molecular transcriptomic alterations [32]. Members of the E2F family are crucial factors in tumorigenesis. Studies have reported that these transcription factors regulate multiple cytokine genes and signaling pathways that are involved in viral tumorigenesis progression [33]. For example, a recent study showed that E2F1 could regulate the expression of DEAD/H-Box helicase 11 (DDX11) to promote tumor progression in HCC [34]. Importantly, E2F4 has been revealed to be a critical regulator of HCC [35].

In this study, we observed that E2F4 expression was dysregulated in pan-cancer, and especially in HCC. Our study has provided a more detailed picture of E2F4 expression in relation to clinical characteristics, prognosis, immune cell infiltration, hub genes, and pathway crosstalk in HCC. Our results demonstrated that E2F4 was highly expressed in HCC, and Kaplan-Meier analyses indicated that E2F4 could be a potential prognostic factor in HCC. Similar to our results, Khaleel and colleagues reported that E2F4 could serve as an outcome prognosis predictor for breast cancer [36]. Another study demonstrated that E2F4 played an essential role in the drug-sensitivity and tumor progression of epithelial ovarian cancer [37]. These findings showed that E2F4 could serve as a novel biomarker for cancer diagnosis and prognosis prediction.

A network analyst algorithm was utilized to further explore the relationship between E2F4 and co-expressed genes, and the crucial enriched

pathways involved in this crosstalk. We found that E2F4 could be identified as a regulator of hub genes in liver cancer. In support of our results, studies have demonstrated that E2F4 function as crucial regulator in liver cancer [38]. The increased E2F4 in breast cancer is also associated with breast cancer liver metastasis [39]. It is also estimated that E2F4 acts as bridging genes in the process of colorectal cancer liver metastasis [40]. Additionally, it was revealed that E2F4 could be a hub node regulatory gene in the subnetwork of HCC. Collectively, these findings suggest that E2F4 may play a critical regulatory role in HCC.

The tumor microenvironment has a strong influence on the carcinogenesis of HCC [41]. The underlying mechanisms of liver immune tolerance are still mostly unknown. Genetic diversity and microenvironment have a close relationship with HCC progression [42]. Therefore, we aimed to expand our current knowledge regarding the E2F4 gene in immune response regulation, taking into consideration the tumor purity and tumor immunity. We found that immune cells have high immune cell purity in HCC. The B cells, CD8<sup>+</sup> T cells, CD4<sup>+</sup> T cells, macrophages, neutrophils, and dendritic cells are co-related immune cells, which serve critical roles in HCC immune infiltration. Altogether, these data powerfully indicate that E2F4 could be a crucial factor mediating immune-associated pathways.

To further investigate the signaling crosstalk involved in controlling abnormal E2F4 gene expression, we examined the E2F4 co-expression network. Our study validates that E2F4 influences cell cycle regulation, DNA repair, and replication. At the same time, E2F4 is also active in metabolic associated process, such as mRNA splicing via the spliceosome and regulation of mitophagy, positive regulation of protein targeting to the mitochondrion, nucleolus, mediator complex ATP binding, nucleotide, binding promoter-specific chromatin binding processes. These molecular activities are consistent with tumor carcinogenesis, and E2F4 could participate in tumor progression via the regulation of these pathways.

Taken together, these findings demonstrate that the up-regulation of E2F4 could be a promising molecular target for the diagnosis and treatment of HCC. Reasonably, the potential role of E2F4 in immune environment regulation and clinical conditional diagnostic function still requires further validation.

## Supplementary Material

Supplementary figure S1.

<http://www.jcancer.org/v12p1792s1.pdf>

## Acknowledgements

We thank the patients and investigators who participated in TCGA for the provided data and the online tools we used for analyses.

## Funding

This study was supported by National Science and Technology Major Project of China (2018ZX10302206 and 2017ZX10202203), and Dr. Zheng was supported by Zhejiang University Academic Award for Outstanding Doctoral Candidates (2020052).

## Author contributions

Zhi Chen, Qiuxian Zheng and Qiang Fu designed the research plan and analysis strategy. Qiuxian Zheng and Jia Xu established the model. Qiuxian Zheng, Xinyu Gu and Haibo Zhou wrote the manuscript and prepared the figures and tables. All authors agreed to be accountable for all aspects of the work and gave final approval of the version to be published.

## Ethics approval and consent to participate

This study was approved by the ethics committee of The First Affiliated Hospital, College of Medicine, Zhejiang University. Written informed consent was obtained from the patients, and the study was conducted in accordance with the Declaration of Helsinki.

## Data availability

All data generated or analyzed during this study are included in this article.

## Competing Interests

The authors have declared that no competing interest exists.

## References

- Cortes J, Perez-García J, Llombart-Cussac A, Curigliano G, El Saghir N, Cardoso F, et al. Enhancing global access to cancer medicines. *CA Cancer J Clin.* 2020; 70: 105-24.
- Wang X, Wang Q, Huang J, Zhang L, Zhou X, Liu J, et al. Novel candidate biomarkers of origin recognition complex 1, 5 and 6 for survival surveillance in patients with hepatocellular carcinoma. *J Cancer.* 2020; 11: 1869-82.
- Craig AJ, von Felden J, Garcia-Lezana T, Sarcognato S, Villanueva A. Tumour evolution in hepatocellular carcinoma. *Nature reviews Gastroenterology & hepatology.* 2020; 17: 139-52.
- Bruix J, da Fonseca LG, Reig M. Insights into the success and failure of systemic therapy for hepatocellular carcinoma. *Nature reviews Gastroenterology & hepatology.* 2019; 16: 617-30.
- Wang X-K, Wang Q-Q, Huang J-L, Zhang L-B, Zhou X, Liu J-Q, et al. Novel candidate biomarkers of origin recognition complex 1, 5 and 6 for survival surveillance in patients with hepatocellular carcinoma. *J Cancer.* 2020; 11: 1869-82.
- Morgunova E, Yin Y, Jolma A, Dave K, Schmierer B, Popov A, et al. Structural insights into the DNA-binding specificity of E2F family transcription factors. *Nature communications.* 2015; 6: 10050.
- Wang H, Wang X, Xu L, Zhang J, Cao H. Integrated analysis of the E2F transcription factors across cancer types. *Oncol Rep.* 2020; 43: 1133-46.
- Zhang K, Zhang B, Bai Y, Dai L. E2F1 promotes cancer cell sensitivity to cisplatin by regulating the cellular DNA damage response through miR-26b in esophageal squamous cell carcinoma. *J Cancer.* 2020; 11: 301-10.
- Parodi S, Ognibene M, Haupt R, Pezzolo A. The Over-Expression of E2F3 Might Serve as Prognostic Marker for Neuroblastoma Patients with Stage 4B Disease. *Diagnostics (Basel).* 2020;10(5):315.
- Huang Y-L, Ning G, Chen L-B, Lian Y-F, Gu Y-R, Wang J-L, et al. Promising diagnostic and prognostic value of E2Fs in human hepatocellular carcinoma. *Cancer Manag Res.* 2019; 11: 1725-40.
- Rodríguez JL, Boukaba A, Sandoval J, Georgieva EI, Latasa MU, García-Trevijano ER, et al. Transcription of the MAT2A gene, coding for methionine adenosyltransferase, is up-regulated by E2F and Sp1 at a chromatin level during proliferation of liver cells. *The international journal of biochemistry & cell biology.* 2007; 39: 842-50.
- Hsu J, Arand J, Chaikovskiy A, Mooney NA, Demeter J, Brison CM, et al. E2F4 regulates transcriptional activation in mouse embryonic stem cells independently of the RB family. *Nature communications.* 2019; 10: 2939.
- Gaubatz S, Lees JA, Lindeman GJ, Livingston DM. E2F4 is exported from the nucleus in a CRM1-dependent manner. *Mol Cell Biol.* 2001; 21: 1384-92.
- Molina-Privado I, Jiménez-P R, Montes-Moreno S, Chiodo Y, Rodríguez-Martínez M, Sánchez-Verde L, et al. E2F4 plays a key role in Burkitt lymphoma tumorigenesis. *Leukemia.* 2012; 26: 2277-85.
- Gong J, Fan H, Deng J, Zhang Q. LncRNA HAND2-AS1 represses cervical cancer progression by interaction with transcription factor E2F4 at the promoter of C16orf74. *J Cell Mol Med.* 2020; 24: 6015-27.
- Paquin M-C, Leblanc C, Lemieux E, Bian B, Rivard N. Functional impact of colorectal cancer-associated mutations in the transcription factor E2F4. *Int J Oncol.* 2013; 43: 2015-22.
- Feng Y, Li L, Du Y, Peng X, Chen F. E2F4 functions as a tumour suppressor in acute myeloid leukaemia via inhibition of the MAPK signalling pathway by binding to EZH2. *J Cell Mol Med.* 2020; 24: 2157-68.
- Zamani-Ahmadmehdi M, Najafi A, Nassiri SM. Reconstruction of canine diffuse large B-cell lymphoma gene regulatory network: detection of functional modules and hub genes. *J Comp Pathol.* 2015; 152: 119-30.
- Xue C, He Y, Zhu W, Chen X, Yu Y, Hu Q, et al. Low expression of LACTB promotes tumor progression and predicts poor prognosis in hepatocellular carcinoma. *American journal of translational research.* 2018; 10: 4152-62.
- Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi BVSK, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. *Neoplasia.* 2017; 19: 649-58.
- Menyhárt O, Nagy Á, Györfy B. Determining consistent prognostic biomarkers of overall survival and vascular invasion in hepatocellular carcinoma. *R Soc Open Sci.* 2018; 5: 181006.
- Sun C-C, Li S-J, Hu W, Zhang J, Zhou Q, Liu C, et al. Comprehensive Analysis of the Expression and Prognosis for E2Fs in Human Breast Cancer. *Molecular therapy : the journal of the American Society of Gene Therapy.* 2019; 27: 1153-65.
- Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer Discov.* 2012; 2: 401-4.
- Li T, Fan J, Wang B, Traugh N, Chen Q, Liu J, et al. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res.* 2017; 77: e108-e10.
- Vasaikar SV, Straub P, Wang J, Zhang B. LinkedOmics: analyzing multi-omics data within and across 32 cancer types. *Nucleic Acids Res.* 2018; 46: D956-D63.
- Tang Z, Kang B, Li C, Chen T, Zhang Z. GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res.* 2019; 47: W556-W60.
- Sun J, Zhou C, Ma Q, Chen W, Atyah M, Yin Y, et al. High GCLC level in tumor tissues is associated with poor prognosis of hepatocellular carcinoma after curative resection. *J Cancer.* 2019; 10: 3333-43.
- Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal.* 2013; 6: p11.
- Gunasekaran G, Bekki Y, Lourdasamy V, Schwartz M. Surgical Treatments of Hepatobiliary Cancers. *Hepatology (Baltimore, Md).* 2020;5:31325.
- Chen Q, Li F, Zhong C, Zou Y, Li Z, Gao Y, et al. Inflammation Score System using Preoperative Inflammatory Markers to Predict Prognosis for Hepatocellular Carcinoma after Hepatectomy: A Cohort Study. *J Cancer.* 2020; 11: 4947-56.
- Rebouissou S, Nault J-C. Advances in molecular classification and precision oncology in hepatocellular carcinoma. *J Hepatol.* 2020; 72: 215-29.
- Gong Z, Yu J, Yang S, Lai PBS, Chen GG. FOX transcription factor family in hepatocellular carcinoma. *Biochimica et biophysica acta Reviews on cancer.* 2020; 1874: 188376.
- Kent LN, Leone G. The broken cycle: E2F dysfunction in cancer. *Nature reviews Cancer.* 2019; 19: 326-38.
- Yu Y, Zhao D, Li K, Cai Y, Xu P, Li R, et al. E2F1 mediated DDX11 transcriptional activation promotes hepatocellular carcinoma progression through PI3K/AKT/mTOR pathway. *Cell Death Dis.* 2020; 11: 273.
- Wang J, Tian Y, Chen H, Li H, Zheng S. Key signaling pathways, genes and transcription factors associated with hepatocellular carcinoma. *Mol Med Report.* 2018; 17: 8153-60.
- Khaleel SS, Andrews EH, Ung M, DiRenzo J, Cheng C. E2F4 regulatory program predicts patient survival prognosis in breast cancer. *Breast cancer research : BCR.* 2014; 16: 486.

37. Long X, Song K, Hu H, Tian Q, Wang W, Dong Q, et al. Long non-coding RNA GAS5 inhibits DDP-resistance and tumor progression of epithelial ovarian cancer via GAS5-E2F4-PARP1-MAPK axis. *Journal of experimental & clinical cancer research* : CR. 2019; 38: 345.
38. Yang W-X, Pan Y-Y, You C-G. CDK1, CCNB1, CDC20, BUB1, MAD2L1, MCM3, BUB1B, MCM2, and RFC4 May Be Potential Therapeutic Targets for Hepatocellular Carcinoma Using Integrated Bioinformatic Analysis. *BioMed research international*. 2019; 2019: 1245072.
39. Wang L, Li J, Liu E, Kinnebrew G, Zhang X, Stover D, et al. Identification of Alternatively-Activated Pathways between Primary Breast Cancer and Liver Metastatic Cancer Using Microarray Data. *Genes (Basel)*. 2019; 10 (10):753.
40. Gao B, Yu T, Xue D, Sun B, Shao Q, Choudhry H, et al. A multidimensional integration analysis reveals potential bridging targets in the process of colorectal cancer liver metastasis. *PLoS One*. 2017; 12: e0178760.
41. Iwata T, Kondo Y, Kimura O, Morosawa T, Fujisaka Y, Umetsu T, et al. PD-L1MDSCs are increased in HCC patients and induced by soluble factor in the tumor microenvironment. *Sci Rep*. 2016; 6: 39296.
42. Li H, Li Y, Lu J-W, Huo X, Gong Z. Liver-specific androgen receptor knockout attenuates early liver tumor development in zebrafish. *Sci Rep*. 2019; 9: 10645.