

Identification of CDC20 As an Immune Infiltration-Correlated Prognostic Biomarker in Hepatocellular Carcinoma

Chen Xiong

Dalian Medical University

Zhihui Wang

Dalian Medical University

Guifu Wang

Dalian Medical University

Chi Zhang

Yangzhou University Affiliated Northern Jiangsu People's Hospital: Northern Jiangsu People's Hospital

Shengjie Jin

Yangzhou University Affiliated Northern Jiangsu People's Hospital: Northern Jiangsu People's Hospital

Guoqing Jiang

Yangzhou University Affiliated Northern Jiangsu People's Hospital: Northern Jiangsu People's Hospital

Dousheng Bai (✉ drbaidousheng@163.com)

Yangzhou University Affiliated Northern Jiangsu People's Hospital: Northern Jiangsu People's Hospital

<https://orcid.org/0000-0001-8032-7959>

Research Article

Keywords: CDC20, hepatocellular carcinoma, E3 ubiquitin-protein ligase, immune infiltration, prognosis

Posted Date: April 14th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-400598/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Hepatocellular carcinoma (HCC) is a malignancy with a poor prognosis. Some E3 ubiquitin-protein ligases play essential roles in HCC development. We aimed to explore a hub E3 ubiquitin-protein ligase gene and verify its association with prognosis and immune cell infiltration in HCC. We identified cell division cycle 20 (CDC20) as a hub E3 ubiquitin-protein ligase in HCC by determining the intersecting genes in a protein-protein interaction (PPI) network of differentially expressed genes (DEGs) in HCC data from the International Cancer Genome Consortium (ICGC) and 919 E3 ubiquitin-protein ligase genes from the Integrated annotations for Ubiquitin and Ubiquitin-like Conjugation Database (IUUCD). DEGs and their correlations with clinicopathological features were explored in The Cancer Genome Atlas (TCGA), ICGC, and Gene Expression Omnibus (GEO) databases via the Wilcoxon signed-rank test. The prognostic value of CDC20 was illustrated by Kaplan-Meier (K-M) curves and Cox regression analyses. Subsequently, the correlation between CDC20 and immune infiltration was demonstrated via the Tumor Immune Estimation Resource (TIMER) and Gene Expression Profiling Interactive Analysis (GEPIA). CDC20 expression was significantly higher in HCC than in normal tissues (all $P < 0.05$). K-M curves and Cox regression analyses showed that high CDC20 expression predicted a poor prognosis and might be an independent risk factor for HCC prognosis ($P < 0.05$). Additionally, the TIMER and GEPIA results indicated that CDC20 is correlated with the immune infiltration of CD8 + T cells, T cells (general), monocytes, and exhausted T cells. This research revealed the potential prognostic value of CDC20 in HCC and demonstrated that CDC20 might be an immune-associated therapeutic target in HCC because of its correlation with immune infiltration.

Introduction

Hepatocellular carcinoma (HCC) is the sixth most common primary tumour and the fourth leading cause of cancer-related death worldwide[1]. Surgical resection is the preferred treatment for HCC, and many new therapeutic strategies, such as targeted therapy and immunotherapy, can be used to treat this disease[2]. However, due to the characteristics of invasion, metastasis, and immune escape, though many advances have improved the treatment of HCC, its prognosis remains dismal[3, 4]. Therefore, it is necessary to explore the relevant mechanism of HCC progression and identify novel therapeutic targets to improve the prognosis of HCC.

E3 ubiquitin-protein ligases are a series of enzymes that catalyse the degradation of proteins by transferring ubiquitin to their substrates[5]. As an increasing number of E3 ubiquitin-protein ligases have been discovered to play oncogenic or anti-oncogenic roles in cancers, the important roles of E3 ubiquitin-protein ligases in the development and carcinogenesis of malignancies have been illustrated. Some E3 ubiquitin ligases are correlated with chemoresistance and may be potential cancer treatment targets[6, 7]. Therefore, it is of great significance to explore hub E3 ubiquitin-protein ligases in HCC to identify new biomarkers and treatment targets.

The immune microenvironment, a crucial part of the tumour microenvironment, is composed of many immune cells, such as T cells, natural killer (NK) cells, and macrophages[8]. It has been demonstrated to have a crucial association with cancer prognosis. For instance, tumours with a regulatory immune microenvironment are usually associated with a terrible prognosis[9]. Recently, immunotherapy has become popular due to its excellent curative effect, which is associated with the components and mechanisms of the immune microenvironment. Immunotherapy targets, such as PD-1, PDL-1, and CTLA4, are biomarkers of relevant immune cells in the immune microenvironment[10, 11]. Thus, it is important to explore novel therapeutic targets associated with the immune microenvironment to identify new treatment strategies for cancer.

In this research, we identified (cell division cycle 20 (CDC20) as a hub E3 ubiquitin-protein ligase in HCC. Subsequently, we explored the prognostic value of CDC20 in HCC by analysing mRNA-sequencing data and clinical prognostic information downloaded from several databases. Then, we analysed the relation between CDC20 and immune cells in the immune environment using an online database. The outcomes illustrated an oncogenic role for CDC20 and a possible association between CDC20 and the tumour immune environment in HCC.

Materials And Methods

Data and information

A total of 919 E3 ubiquitin-protein ligases were obtained from the Integrated annotations for Ubiquitin and Ubiquitin-like Conjugation Database (IUUCD) (<http://iuucd.biocuckoo.org/>), a database that contains information on different ubiquitinating and deubiquitinating enzymes[12]. The gene expression data and corresponding clinical information of 231 HCC patients were acquired from the International Cancer Genome Consortium (ICGC) database (<https://dcc.icgc.org/projects/LIRI-JP>)[13]. All patients in the ICGC cohort were from Japan. Other RNA-sequencing data on CDC20 and the prognostic data on an additional 371 HCC patients were obtained from The Cancer Genome Atlas (TCGA) database up to January 3, 2021 (<https://portal.gdc.cancer.gov/>)[14]. In addition, other expression data on CDC20 and the clinicopathologic features of HCC patients were downloaded from the “GSE14520”, “GSE45267”, “GSE76427”, and “GSE121248” datasets in the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>)[15]. Immunohistochemical figures were obtained from the Human Protein Atlas (HPA) database (<https://www.proteinatlas.org/>).

Identifying hub E3 ubiquitin-protein ligases

We explored potential hub E3 ubiquitin-protein ligases by taking the intersecting genes of 919 E3 ubiquitin-protein ligases and differentially expressed genes (DEGs) in the ICGC ($\log FC > 2$). The protein-protein interaction (PPI) network of the intersecting genes was analysed with STRING (<https://www.string-db.org/>)[16]. The degrees of each intersecting gene in the PPI network were calculated with Cytoscape software (<https://cytoscape.org/>)[17]. The gene with the most degrees was identified as the hub E3 ubiquitin-protein ligase in HCC.

Differential expression analysis

CDC20 was identified as the hub E3 ubiquitin-protein ligase in HCC. First, datasets from the TCGA, ICGC and GEO databases (“GSE14520”, “GSE45267”, “GSE76427”, and “GSE121248”) were used to analyse the expression levels of CDC20 in HCC samples and normal hepatic samples. For the analysis, the Wilcoxon signed-rank test and the “limma” package in R software 4.0.0 (<https://www.r-project.org/>) were utilized. The results were drawn with the “beeswarm” package in R software. In addition, we acquired immunohistochemical images of CDC20 in HCC tissues and normal hepatic tissues from the HPA database.

Evaluating the prognostic value of CDC20

Based on the information on CDC20 expression and the overall survival of HCC patients obtained from the TCGA, ICGC, and GEO (GSE14520) databases, Kaplan-Meier (K-M) curves and scatterplots were plotted with the “survival” and “pheatmap” packages in R software to show the relationship between CDC20 and the prognosis of HCC. Additionally, receiver operating characteristic (ROC) curves were drawn to determine the areas under the curve (AUCs) of CDC20 at 1 year, 2 years, and 3 years, which were plotted via the “survival”, “survminer”, and “timeROC” packages in R software.

Analysing the expression of CDC20 according to clinical features

To further explore the prognostic value of CDC20, we analysed the expression of CDC20 according to clinical features via univariate and multivariate Cox regression analyses with the “survival” package in R software. The AUCs of the ROC curves were used to compare the prognostic values of CDC20 and clinical features via the “survivalROC” package. Then, we investigated the correlations between the expression level of CDC20 and the clinical characteristics of patients with the “ggpubr” package in R software and Perl language (<https://www.perl.org/>). All data and information were obtained from the TCGA database.

Gene set enrichment analysis (GSEA) of CDC20

We divided all HCC samples in the TCGA database into the CDC20^{low} group and the CDC20^{high} group according to the median expression level of CDC20. Single-gene GSEA was employed to explore the significant biological pathways between the two groups with GSEA_4.0.1 software (<https://www.gsea-msigdb.org/gsea/index.jsp>). All results were plotted via the “ggplot2” package.

Immune infiltration analysis of CDC20

The Tumor Immune Estimation Resource (TIMER) (<https://cistrome.shinyapps.io/timer/>) is a website used to systematically analyse the correlation between gene expression and the infiltration of immune cells with data on different cancers in the TCGA[18]. We utilized TIMER to analyse the associations between CDC20 and 6 kinds of immune cells (CD4 + T cells, CD8 + T cells, B cells, macrophages, dendritic cells (DC), and neutrophils) in the tumour microenvironment of HCC through a gene module. Corresponding results were plotted with TIMER to demonstrate correlations between the CDC20

expression level and tumour purity, a crucial element that can affect the analysis of immune infiltration via genomic approaches[19, 20]. The size of the partial correlation coefficient (partial cor) was utilized to reveal the degree of association between the expression of CDC20 and immune cell infiltration. Subsequently, multivariate Cox regression analysis was performed to analyse the prognostic values of CDC20 and these 6 types of immune cells via the survival module in TIMER. The correlation between CDC20 and biomarkers of tumour-infiltrating immune cells (TIICs) was also explored through the “correlation” module. These immune cells consisted of CD8 + T cells, T cells (general), B cells, monocytes, tumour-associated macrophages (TAMs), M1 macrophages, M2 macrophages, neutrophils, T-helper 1 (Th1) cells, T-helper 2 (Th2) cells), T-helper 17 (Th17) cells), follicular helper T (Tfh) cells), regulatory T cells (Tregs), NK cells, DCs, and exhausted T cells. All of these immune cell types have been previously described[19, 21–24]. Gene Expression Profiling Interactive Analysis (GEPIA) (<http://gepia.cancer-pku.cn>) is an online database that contains RNA-sequencing data on 369 HCC samples and 160 normal liver samples from the TCGA and Genotype-Tissue Expression (GTEx) project[25]. By using GEPIA via Spearman’s correlation, we further illustrated the associations between the expression of CDC20 and the gene biomarkers of four types of immune cells that had the most significant correlations with CDC20 according to the TIMER results. Additionally, the value of R was employed to explain the correlation coefficient between CDC20 and these 4 types of immune cells according to the GEPIA results.

Statistical analysis

All statistical analytical methods were based on the Perl language and R software 4.0.0. The expression data of CDC20 from 3 different databases were normalized by log₂ transformation. K-M survival curves were calculated based on log-rank tests. The associations of CDC20 with immune infiltration and gene markers of different immune cells were calculated by Spearman’s correlation. The associations of immune infiltration were determined via the following guide for the partial cor value: 0.00–0.19 “very weak,” 0.20–0.39 “weak,” 0.40–0.59 “moderate,” 0.60–0.79 “strong,” and 0.80–1.0 “very strong”[26]. All P values less than 0.05 were considered to indicate statistical significance.

Results

CDC20 was identified as the hub gene

The entire research process is plotted in Fig. 1. A total of 859 DEGs obtained from the ICGC (log FC > 2) are shown in a volcano map (Fig. 2a). In total, 21 intersecting genes were identified (intersecting genes between DEGs in the ICGC and 919 E3 ubiquitin-protein ligases) (Fig. 2b). Among these 21 genes, the PPI network of 16 genes with degrees > 0 is shown in Fig. 2c. From this figure, we identified CDC20 as the hub gene with the most connections between these 21 genes.

The expression level of CDC20 was higher in HCC samples than in normal samples To analyse the differential expression of CDC20 between HCC and normal samples, expression data from the TCGA, ICGC, and GEO databases were analysed. The mRNA levels of CDC20 from the three cohorts are shown in Fig. 3. The immunohistochemical staining results of CDC20 in HCC from the HPA database are also

shown in Fig. 3. According to this figure, it can be speculated that the expression level of CDC20 in all 4 databases was remarkably lower in normal hepatic samples than in HCC samples ($P < 0.05$).

CDC20 predicts a poor prognosis in HCC

After exploring the expression of CDC20 in HCC, we verified the prognostic value of CDC20 in HCC by survival analysis via K-M curves. Patients were ranked in ascending order by the expression level of CDC20 and categorized into a high-expression group and a low-expression group according to the median expression level of CDC20. The K-M curves of CDC20 are plotted in Fig. 4a-c and show the prognostic value of CDC20 according to data obtained from the TCGA, ICGC, and GEO datasets. Survival analysis revealed that the prognosis of patients in all three cohorts with low CDC20 expression was much better than that of patients with high CDC20 expression ($P < 0.05$). All patients' survival times and survival conditions are shown in scatterplots (Fig. 4d-f). Patients in the CDC20^{low} group had a remarkably longer survival time and a lower survival rate than those in the CDC20^{high} group. Additionally, the ROC curves demonstrated the ability of CDC20 to predict overall survival at 1 year, 2 years, and 3 years based on data from 3 datasets (Fig. 4g-i). These data demonstrate that CDC20 predicts a poor prognosis in HCC patients.

Correlations between CDC20 and clinical features

We analysed the correlations between the expression of CDC20 and clinical features via TCGA data. The statistically significant results are shown in Fig. 5a-c. According to this figure, it can be speculated that the expression of CDC20 is associated with pathological grade (Fig. 5a), TNM stage (Fig. 5b), and T stage (Fig. 5c). CDC20 was highly expressed in samples with a higher pathological grade, TNM stage, and T stage (all $P < 0.01$), indicating that CDC20 plays an oncogenic role in HCC.

Comparing the prognostic values of CDC20 and clinical features

To further validate the prognostic value of CDC20, we compared that of CDC20 with that of other clinical features. First, univariate and multivariate Cox regression analyses were performed (Fig. 5d,e). All the results showed that CDC20 had independent predictive value ($P < 0.05$). Moreover, CDC20 had a higher AUC value (AUC = 0.728) than the other clinical features, which indicates that CDC20 has more prognostic value in HCC (Fig. 5f).

GSEA of CDC20

The potential biological functions of CDC20 in HCC were preliminarily explored via GSEA. The GSEA results are shown in Fig. 6. High expression of CDC20 might be involved in several signalling pathways related to cancers, such as the P53 signalling pathway, mTOR signalling pathway, and NOTCH signalling pathway. The high expression of CDC20 was also associated with the cell cycle, ubiquitin-mediated proteolysis, and T cell receptor signalling pathways.

Association between CDC20 and immune cell infiltration in HCC

GSEA showed some pathways associated with CDC20, some of which were demonstrated to be correlated with the immune microenvironment and the immune response[27, 28]. Thus, we investigated the association between CDC20 and 6 different immune cell types via TIMER. We observed that the level of CDC20 expression positively correlated with the infiltration levels of B cells (partial cor = 0.453, $p = 8.39E-19$), CD8 + T cells (partial cor = 0.352, $p = 2.01E-11$), CD4 + T cells (partial cor = 0.267, $p = 4.78E-07$), macrophages (partial cor = 0.419, $p = 6.63E-16$), neutrophils (partial cor = 0.315, $p = 2.12E-09$), and DCs (partial cor = 0.46, $p = 3.54E-19$) (Fig. 7a), suggesting an association between the expression of CDC20 and immune cell infiltration in HCC. We also found an association between CDC20 and tumour purity (partial cor = 0.129, $p = 1.67E-02$) in HCC. To explore the prognostic value of CDC20 in HCC, multivariate Cox regression analysis of CDC20 and 6 types of immune cells was performed via TIMER (Fig. 7b). The analysis revealed that CDC20 had more independent prognostic value than the 6 types of immune cells ($P < 0.05$).

Correlations between CDC20 and gene markers of immune cells in HCC

The immune cell infiltration analysis suggested an association between CDC20 and immune cells in HCC. To further investigate why CDC20 correlated with high immune cell infiltration, the correlations between CDC20 and gene markers of several types of immune cells were analysed via TIMER and GEPIA.

The TIMER results are listed in Table 1. The expression of CDC20 was remarkably correlated with the gene biomarkers of CD8 + T cells (CD8A and CD8B) after adjusting for tumour purity, as shown in Table 1 and Fig. 8a (all $P < 0.05$), which indicates that CDC20 is related to a specific immune response. CDC20 was also demonstrated to be associated with gene biomarkers of other two other immune cell types, namely, T cells (general) (CDC3D, CDC3E, and CD2) and monocytes (CD86 and CSF1R), after adjusting for tumour purity (Table 1 and Fig. 8b,c) (all $P < 0.05$). In addition, the gene biomarkers of exhausted T cells, including CTLA4, PD-1, LAG3, GZMB, and TIM-3, were illustrated to be correlated with the expression of CDC20 in HCC (Table 1 and Fig. 8d) ($P < 0.05$). Some gene biomarkers, such as PD-1 and CTLA4, have proven to be targets of immunotherapy, suggesting an association between CDC20 and immunotherapy in HCC. Subsequently, we validated the association between CDC20 and the gene biomarkers of 4 types of immune cells through the GEPIA website. All results were significantly different (all $P < 0.05$) (Table 2).

Table 1
Correlation analysis between CDC20 and relative gene biomarkers of immune cells in TIMER.

Description	Gene markers	HCC			
		None		Purity	
		partial cor	<i>p</i> value	partial cor	<i>p</i> value
CD8 + T cell	CD8A	0.218	***	0.306	***
	CD8B	0.256	***	0.34	***
T cell (general)	CD3D	0.364	***	0.474	***
	CD3E	0.217	***	0.354	***
	CD2	0.243	***	0.367	***
B cell	CD19	0.275	***	0.342	***
	CD79A	0.165	*	0.26	***
Monocyte	CD86	0.328	***	0.462	***
	CD115(CSF1R)	0.179	**	0.301	***
TAM	CCL2	0.038	4.64E-01	0.116	3.19E-02
	CD68	0.242	***	0.319	***
	IL10	0.23	***	0.331	***
M1 Macrophage	iNOS(NOS2)	-0.098	5.88E-02	-0.09	9.64E-02
	IRF5	0.325	***	0.332	***
	COX2(PTGS2)	0.034	5.14E-01	0.128	1.70E-02
M2 Macrophage	CD163	0.046	3.8E-01	0.132	1.43E-02
	VSIG4	0.098	5.88E-02	0.189	**
	MS4A4A	0.089	8.69E-02	0.199	**

Abbreviation: HCC: hepatocellular carcinoma; TAM: tumor-associated macrophage; Th1, T-helper 1 cells; Th2, T-helper 2 cells; Th17, T-helper 17 cells; Tfh: Follicular helper T cell; Treg, regulatory T cell. Note: Partial cor (partial correlation coefficient) represents the correlation coefficient between the CDC20 expression and gene markers of infiltrating immune cells; Purity represents the correlation adjusted by purity; None represents the correlation without adjustment; * $p < 0.01$; ** $p < 0.001$; *** $p < 0.0001$.

Description	Gene markers	HCC			
		None		Purity	
		partial cor	<i>p</i> value	partial cor	<i>p</i> value
Neutrophils	CD66b(CEACAM8)	0.092	7.75E-02	0.118	2.84E-02
	CD11b(ITGAM)	0.317	***	0.403	***
	CCR7	0.035	4.99E-01	0.144	*
Natural killer cell	KIR2DL1	-0.058	2.62E-01	-0.082	1.26E-01
	KIR2DL3	0.139	*	0.173	*
	KIR2DL4	0.228	***	0.245	***
	KIR3DL1	-0.033	5.28E-01	-0.026	6.32E-01
	KIR3DL2	0.095	0.0662	0.129	0.0162
	KIR3DL3	0.062	2.31E-01	0.07	1.97E-01
	KIR2DS4	0.018	7.3E-01	0.012	8.26E-01
	Dendritic cell	HLA-DPB1	0.196	**	0.297
	HLA-DQB1	0.193	**	0.281	***
	HLA-DRA	0.18	**	0.28	***
	HLA-DPA1	0.151	*	0.256	***
	BDCA-1(CD1C)	0.063	2.25E-01	0.137	1.09E-02
	BDCA-4(NRP1)	0.114	2.77E-02	0.13	1.59E-02
	CD11c (ITGAX)	0.318	***	0.433	***
Th1	T-bet (TBX21)	0.073	1.59E-01	0.162	*

Abbreviation: HCC: hepatocellular carcinoma; TAM: tumor-associated macrophage; Th1, T-helper 1 cells; Th2, T-helper 2 cells; Th17, T-helper 17 cells; Tfh: Follicular helper T cell; Treg, regulatory T cell. Note: Partial cor (partial correlation coefficient) represents the correlation coefficient between the CDC20 expression and gene markers of infiltrating immune cells; Purity represents the correlation adjusted by purity; None represents the correlation without adjustment; * $p < 0.01$; ** $p < 0.001$; *** $p < 0.0001$.

Description	Gene markers	HCC			
		None		Purity	
		partial cor	<i>p</i> value	partial cor	<i>p</i> value
	STAT4	0.227	***	0.28	***
	STAT1	0.287	***	0.32	***
	IFN- γ (IFNA1)	0.056	2.84E-01	0.062	2.54E-01
	TNF- α	0.265	***	0.378	***
Th2	GATA3	0.203	***	0.316	***
	STAT6	-0.026	6.2E-01	-0.032	5.58E-01
	STAT5A	0.26	***	0.299	***
	IL13	0.101	5.2E-02	0.109	4.24E-02
Tfh	BCL6	0.075	1.5E-01	0.086	1.10E-01
	IL21	0.136	*	0.166	*
Th17	STAT3	0.166	*	0.056	3.01E-01
	IL17A	0.043	4.11E-01	0.066	2.24E-01
Treg	FOXP3	0.097	6.12E-02	0.174	*
	CCR8	0.31	***	0.39	***
	STAT5B	0.069	1.83E-01	0.06	2.65E-01
	TGF β (TGFB1)	0.27	***	0.354	***
Exhausted T cells	PD-1(PDCD-1)	0.37	***	0.461	***
	CTLA4	0.396	***	0.494	***
	LAG3	0.372	***	0.406	***
	TIM3(HAVCR2)	0.352	***	0.493	***

Abbreviation: HCC: hepatocellular carcinoma; TAM: tumor-associated macrophage; Th1, T-helper 1 cells; Th2, T-helper 2 cells; Th17, T-helper 17 cells; Tfh: Follicular helper T cell; Treg, regulatory T cell. Note: Partial cor (partial correlation coefficient) represents the correlation coefficient between the CDC20 expression and gene markers of infiltrating immune cells; Purity represents the correlation adjusted by purity; None represents the correlation without adjustment; * $p < 0.01$; ** $p < 0.001$; *** $p < 0.0001$.

Description	Gene markers	HCC			
		None		Purity	
		partial cor	<i>p</i> value	partial cor	<i>p</i> value
	GZMB	0.121	1.96E-02	0.182	**

Abbreviation: HCC: hepatocellular carcinoma; TAM: tumor-associated macrophage; Th1, T-helper 1 cells; Th2, T-helper 2 cells; Th17, T-helper 17 cells; Tfh: Follicular helper T cell; Treg, regulatory T cell. Note: Partial cor (partial correlation coefficient) represents the correlation coefficient between the CDC20 expression and gene markers of infiltrating immune cells; Purity represents the correlation adjusted by purity; None represents the correlation without adjustment; * $p < 0.01$; ** $p < 0.001$; *** $p < 0.0001$.

Table 2
Correlation analysis between CDC20 and relative gene biomarkers of immune cells in GEPIA.

Description	Gene markers	HCC			
		Tumor		Normal	
		R	<i>p</i> value	R	<i>p</i> value
CD8 + T cell	CD8A	0.18	0.00041	0.32	0.024
	CD8B	0.24	3.6E-06	0.26	0.064
T cell (general)	CD3D	0.4	3.1E-15	0.34	0.017
	CD3E	0.19	0.00022	0.33	0.02
	CD2	0.23	8.4E-06	0.32	0.023
Monocyte	CD86	0.27	1.4E-07	0.14	0.35
	CD115(CSF1R)	0.11	0.029	0.18	0.21
Exhausted T cell	PD-1(PDCD1)	0.35	3.1E-12	0.37	0.0081
	CTLA4	0.39	8.8E-15	0.27	0.057
	LAG3	0.36	1.1E-12	0.15	0.29
	TIM-3(HAVCR2)	0.3	2.6E-09	0.04	0.78
	GZMB	0.12	0.025	0.33	0.021

Discussion

HCC is a worldwide problem because of its high mortality and morbidity. It is crucial to identify new targets for the diagnosis and therapy of HCC. Many E3 ubiquitin-protein ligases have been illustrated to play essential roles and might be potential therapy targets in several carcinomas, including HCC[5, 29–31]. Immunotherapy is a widely used treatment for many malignancies, including HCC. Immune cell infiltration in the immune microenvironment was demonstrated to be associated with immunosurveillance and immunotherapy in HCC[32]. Ubiquitin signalling was demonstrated to be associated with the immune response[33]. The E3 ubiquitin-protein ligase MDM2 was verified to sustain STAT5 stability to control T cell immunity in tumours[34]. However, the association between E3 ubiquitin-protein ligases and immune cell infiltration has not been explored in HCC. Therefore, it was novel and meaningful to identify a hub E3 ubiquitin-protein ligase and analyse its association with immune cell infiltration in HCC.

In this research, we identified CDC20 as a hub E3 ubiquitin-protein ligase in HCC and demonstrated its prognostic value in HCC through data from the TCGA, ICGC, GEO, and TIMER databases. All results showed a strong oncogenic function of CDC20, and the prognosis of HCC patients with high CDC20 expression was poor.

CDC20 is a hub gene in cell cycle progression. CDC20 contributes to the progression or metastasis of several carcinomas, such as pancreatic cancer and breast cancer[35]. Alfarsi et al verified that the expression of CDC20 in oestrogen receptor-positive breast cancer predicted a poor prognosis and no response to endocrine therapy[36]. The high expression of CDC20 cancer also predicts a poor prognosis in patients with gastric cancer, prostate cancer, bladder cancer, and colorectal cancer[37–40]. It has been verified that Cdc20 can induce the radioresistance of bladder cancer by degrading FoxO1[41]. Decreasing the expression level of CDC20 might improve radiosensitivity through Mcl-1/p-Chk1-mediated DNA damage and apoptosis in colorectal cancer[42]. In HCC, Shi et al[43] demonstrated that CDC20 could promote PHD3 ubiquitination and activate HIF-1 signalling to accelerate the proliferation of cancer cells. Additionally, Li et al[44] illustrated that the increased expression of CDC20 was associated with HCC proliferation. We performed GSEA to explore potential biological functions and signalling pathways associated with CDC20 in HCC. The results showed that the major functions and pathways were enriched in the P53 signalling, mTOR signalling, NOTCH signalling, cell cycle, ubiquitin-mediated proteolysis, and T cell receptor signalling pathways. Among these pathways, the P53, mTOR and NOTCH signalling pathways are correlated with HCC metastasis, which is the main reason for a poor prognosis[45, 46]. Furthermore, the NOTCH and T cell receptor signalling pathways are thought to be involved in some cellular elements of the tumour microenvironment[47]. Therefore, the possible associations between CDC20 and certain malignant phenotypes, including metastasis and immune infiltration, should be explored and verified.

All previous studies explored only the prognostic value and oncogenic function of CDC20 in cancers. Nevertheless, the possible association between CDC20 and immune cell infiltration in the tumour environment had not been explored. In this research, we analysed the correlation between CDC20 and

immune cell infiltration in HCC and illustrated a possible relation between CDC20 and the immune microenvironment of HCC.

At the beginning of tumour metastasis, tumour cells escape from the anti-tumour immune response and adapt to the microenvironment of the next site of metastasis[48]. Many immune cells in the immune microenvironment play crucial roles in the process of tumour metastasis. For instance, TAMs and neutrophils are promote tumour proliferation and metastasis[49]. Targeting TAMs might be a potential treatment strategy for HCC[50]. In our research, we surveyed the possible correlations between CDC20 expression and immune cell infiltration in HCC. The results suggested significant connections between CDC20 and the infiltration of several immune cell types, including B cells, CD4 + T cells, CD8 + T cells, neutrophils, macrophages, and DCs. In addition, the correlations between CDC20 and relative gene biomarkers of 4 types of immune cells were verified through TIMER and GEPIA. These gene markers include several special genes involved in immunotherapy, such as PDCD1 and CTLA4[51]. The correlation analysis indicated that the expression level of CDC20 influences the effect of immunotherapy, which hinted CDC20 as a potential therapeutic target correlated with tumor immunology.

Conclusions

In summary, a high expression level of CDC20 in HCC indicates a poor prognosis. Additionally, high CDC20 expression is remarkably correlated with immune infiltration. The correlations between CDC20 and gene biomarkers of several immune cell types indicate a role for CDC20 in the immunology of HCC. Further experiments are supposed to test to explore the potential biological mechanisms of CDC20 involved in the metastasis and immune infiltration of HCC. All results showed that CDC20 might be an innovative prognostic biomarker and potential therapeutic target in HCC.

Abbreviations

HCC
hepatocellular carcinoma; CDC20:cell division cycle 20; IUUCD:Integrated annotations for Ubiquitin and Ubiquitin-like Conjugation Database; ICGC:International Cancer Genome Consortium; TCGA:The Cancer Genome Atlas; GEO:Gene Expression Omnibus; HPA:Human Protein Atlas; DEGs:differentially expressed genes; E3 genes:E3 ubiquitin-protein ligase genes; K-M:Kaplan-Meier; ROC:receiver operating characteristic; AUC:the area under the ROC curve; GSEA:gene set enrichment analysis; PPI:protein-protein interaction; TIMER:Tumor Immune Estimation Resource; TIICs:tumor-infiltrating immune cells; TAMs:tumor-associated macrophages; Th1:T-helper 1 cells; Th2:T-helper 2 cells; Th17:T-helper 17 cells; Tfh:follicular helper T cells; Tregs:Regulatory T cells; NK cells:natural killer cells; DCs:dendritic cells; GEPIA:Gene Expression Profiling Interactive Analysis; TNM:Tumor size/lymph nodes/distance metastasis.

Declarations

Funding

This work was supported by the National Natural Science Foundation of China [Grant No.81871909]; “13th five-year Plan” Science and Education strong Health Project leading personnel of Yangzhou [Grant No.LJRC20181]; and Provincial-level discipline leader of the NJPH [Grant No.DTRC201809].

Conflicts of interest

All authors declare that they have no competing interests.

Availability of data and materials

The datasets analyzed were downloaded from The Cancer Genome Atlas (TCGA) database (<https://portal.gdc.cancer.gov/>), International Cancer Genome Consortium (ICGC) database (<https://dcc.icgc.org/projects/LIRI-JP>), Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo>), and Integrated annotations for Ubiquitin and Ubiquitin-like Conjugation Database (IUUCD) database (<http://iuucd.biocuckoo.org/>).

Code availability

Not applicable.

Authors' contributions

CX, ZW and GW contributed to the conception and design of the study; CX collected data and wrote the manuscript; ZW and GW accomplished the data analysis and drew the figures and tables; CX and ZW discussed the results of this research. CX, CZ, SJ, GJ, and DB reviewed and altered the manuscript of the study. In addition, DB was responsible for the organization, revision and submission of this manuscript. All authors read and approved the final manuscript.

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

All authors consent to the publication of this research.

Acknowledgments

The authors greatly thank Dingwen Wang (University of Waikato, Hamilton, New Zealand) and Yahuan Xie (Jiangxi Qiande Culture Communication Co, Ltd, Nanchang, China) for helping acquire the data from

databases in this research. Besides, Chen Xiong especially feel thankful to his parents for their encouragements and supports during the last twenty years.

References

1. Villanueva A (2019) Hepatocellular Carcinoma. *The New England journal of medicine* 380(15):1450-1462.
2. Llovet JM, Montal R, Sia D, Finn RS (2018) Molecular therapies and precision medicine for hepatocellular carcinoma. *Nature reviews Clinical oncology* 15(10):599-616.
3. Li R, Wang Y, Zhang X, Feng M, Ma J, Li J, Yang X, Fang F, Xia Q, Zhang Z et al (2019) Exosome-mediated secretion of LOXL4 promotes hepatocellular carcinoma cell invasion and metastasis. *Molecular cancer* 18(1):18.
4. Li W, Wang H, Ma Z, Zhang J, Ou-Yang W, Qi Y, Liu J (2019) Multi-omics Analysis of Microenvironment Characteristics and Immune Escape Mechanisms of Hepatocellular Carcinoma. *Frontiers in oncology* 9:1019.
5. Wang D, Ma L, Wang B, Liu J, Wei W (2017) E3 ubiquitin ligases in cancer and implications for therapies. *Cancer metastasis reviews* 36(4):683-702.
6. Liu J, Peng Y, Zhang J, Long J, Liu J, Wei W (2020) Targeting SCF E3 Ligases for Cancer Therapies. *Advances in experimental medicine and biology* 1217:123-146.
7. Yang L, Chen J, Huang X, Zhang E, He J, Cai Z (2018) Novel Insights Into E3 Ubiquitin Ligase in Cancer Chemoresistance. *The American journal of the medical sciences* 355(4):368-376.
8. Kurebayashi Y, Ojima H, Tsujikawa H, Kubota N, Maehara J, Abe Y, Kitago M, Shinoda M, Kitagawa Y, Sakamoto M (2018) Landscape of immune microenvironment in hepatocellular carcinoma and its additional impact on histological and molecular classification. *Hepatology (Baltimore, Md)* 68(3):1025-1041.
9. Palucka AK, Coussens LM (2016) The Basis of Oncoimmunology. *Cell* 164(6):1233-1247.
10. Gibney GT, Weiner LM, Atkins MB (2016) Predictive biomarkers for checkpoint inhibitor-based immunotherapy. *The Lancet Oncology* 17(12):e542-e551.
11. Patel SP, Kurzrock R (2015) PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy. *Molecular cancer therapeutics* 14(4):847-856.
12. Zhou J, Xu Y, Lin S, Guo Y, Deng W, Zhang Y, Guo A, Xue Y (2018) iUUCD 2.0: an update with rich annotations for ubiquitin and ubiquitin-like conjugations. *Nucleic acids research* 46(D1):D447-d453.
13. Hudson TJ, Anderson W, Artez A, Barker AD, Bell C, Bernabé RR, Bhan MK, Calvo F, Eerola I, Gerhard DS et al (2010) International network of cancer genome projects. *Nature* 464(7291):993-998.
14. Tomczak K, Czerwińska P, Wiznerowicz M (2015) The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemporary oncology (Poznan, Poland)* 19(1a):A68-77.
15. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M et al (2013) NCBI GEO: archive for functional genomics data sets—update. *Nucleic acids research* 41(Database issue):D991-995.

16. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P et al (2017) The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic acids research* 45(D1):D362-d368.
17. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003) Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 13(11):2498-2504.
18. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B, Liu XS (2017) TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer research* 77(21):e108-e110.
19. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA et al (2013) Inferring tumour purity and stromal and immune cell admixture from expression data. *Nature communications* 4:2612.
20. Aran D, Sirota M, Butte AJ (2015) Systematic pan-cancer analysis of tumour purity. *Nature communications* 6:8971.
21. Danaher P, Warren S, Dennis L, D'Amico L, White A, Disis ML, Geller MA, Odunsi K, Beechem J, Fling SP (2017) Gene expression markers of Tumor Infiltrating Leukocytes. *Journal for immunotherapy of cancer* 5:18.
22. Chen B, Khodadoust MS, Liu CL, Newman AM, Alizadeh AA (2018) Profiling Tumor Infiltrating Immune Cells with CIBERSORT. *Methods in molecular biology (Clifton, NJ)* 1711:243-259.
23. Siemers NO, Holloway JL, Chang H, Chasalow SD, Ross-MacDonald PB, Voliva CF, Szustakowski JD (2017) Genome-wide association analysis identifies genetic correlates of immune infiltrates in solid tumors. *PloS one* 12(7):e0179726.
24. Sousa S, Määttä J (2016) The role of tumour-associated macrophages in bone metastasis. *Journal of bone oncology* 5(3):135-138.
25. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z (2017) GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic acids research* 45(W1):W98-w102.
26. Pan JH, Zhou H, Cooper L, Huang JL, Zhu SB, Zhao XX, Ding H, Pan YL, Rong L (2019) LAYN Is a Prognostic Biomarker and Correlated With Immune Infiltrates in Gastric and Colon Cancers. *Frontiers in immunology* 10:6.
27. Soave DF, Miguel MP, Tomé FD, de Menezes LB, Nagib PR, Celes MR (2016) The Fate of the Tumor in the Hands of Microenvironment: Role of TAMs and mTOR Pathway. *Mediators of inflammation* 2016:8910520.
28. Kreileder M, Barrett I, Bendtsen C, Brennan D, Kolch W (2021) Signaling Dynamics Regulating Crosstalks between T-Cell Activation and Immune Checkpoints. *Trends in cell biology* 31(3):224-235.
29. Zhu L, Qin C, Li T, Ma X, Qiu Y, Lin Y, Ma D, Qin Z, Sun C, Shen X et al (2020) The E3 ubiquitin ligase TRIM7 suppressed hepatocellular carcinoma progression by directly targeting Src protein. *Cell death and differentiation* 27(6):1819-1831.
30. He H, Wu Z, Li S, Chen K, Wang D, Zou H, Chen H, Li Y, Liu Z, Qu C (2020) TRAF7 enhances ubiquitin-degradation of KLF4 to promote hepatocellular carcinoma progression. *Cancer letters* 469:380-389.

31. Shao G, Wang R, Sun A, Wei J, Peng K, Dai Q, Yang W, Lin Q (2018) The E3 ubiquitin ligase NEDD4 mediates cell migration signaling of EGFR in lung cancer cells. *Molecular cancer* 17(1):24.
32. Dong LQ, Peng LH, Ma LJ, Liu DB, Zhang S, Luo SZ, Rao JH, Zhu HW, Yang SX, Xi SJ et al (2020) Heterogeneous immunogenomic features and distinct escape mechanisms in multifocal hepatocellular carcinoma. *Journal of hepatology* 72(5):896-908.
33. Hu H, Sun SC (2016) Ubiquitin signaling in immune responses. *Cell research* 26(4):457-483.
34. Zhou J, Kryczek I, Li S, Li X, Aguilar A, Wei S, Grove S, Vatan L, Yu J, Yan Y et al (2021) The ubiquitin ligase MDM2 sustains STAT5 stability to control T cell-mediated antitumor immunity. *Nature immunology* 22(4):460-470.
35. Cheng S, Castillo V, Sliva D (2019) CDC20 associated with cancer metastasis and novel mushroom-derived CDC20 inhibitors with antimetastatic activity. *International journal of oncology* 54(6):2250-2256.
36. Alfarsi LH, Ansari RE, Craze ML, Toss MS, Masisi B, Ellis IO, Rakha EA, Green AR (2019) CDC20 expression in oestrogen receptor positive breast cancer predicts poor prognosis and lack of response to endocrine therapy. *Breast cancer research and treatment* 178(3):535-544.
37. Mao Y, Li K, Lu L, Si-Tu J, Lu M, Gao X: Overexpression of Cdc20 in clinically localized prostate cancer (2016) Relation to high Gleason score and biochemical recurrence after laparoscopic radical prostatectomy. *Cancer biomarkers : section A of Disease markers* 16(3):351-358.
38. Choi JW, Kim Y, Lee JH, Kim YS (2013) High expression of spindle assembly checkpoint proteins CDC20 and MAD2 is associated with poor prognosis in urothelial bladder cancer. *Virchows Archiv : an international journal of pathology* 463(5):681-687.
39. Ding ZY, Wu HR, Zhang JM, Huang GR, Ji DD (2014) Expression characteristics of CDC20 in gastric cancer and its correlation with poor prognosis. *International journal of clinical and experimental pathology* 7(2):722-727.
40. Li J, Wang Y, Wang X, Yang Q (2020) CDK1 and CDC20 overexpression in patients with colorectal cancer are associated with poor prognosis: evidence from integrated bioinformatics analysis. *World journal of surgical oncology* 18(1):50.
41. Wang L, Yang C, Chu M, Wang ZW, Xue B (2021) Cdc20 induces the radioresistance of bladder cancer cells by targeting FoxO1 degradation. *Cancer letters* 500:172-181.
42. Gao Y, Wen P, Chen B, Hu G, Wu L, Xu A, Zhao G (2020) Downregulation of CDC20 Increases Radiosensitivity through Mcl-1/p-Chk1-Mediated DNA Damage and Apoptosis in Tumor Cells. *International journal of molecular sciences* 21(18):6692.
43. Shi M, Dai WQ, Jia RR, Zhang QH, Wei J, Wang YG, Xiang SH, Liu B, Xu L (2021) APC(CDC20)-mediated degradation of PHD3 stabilizes HIF-1a and promotes tumorigenesis in hepatocellular carcinoma. *Cancer letters* 496:144-155.
44. Li J, Gao JZ, Du JL, Huang ZX, Wei LX (2014) Increased CDC20 expression is associated with development and progression of hepatocellular carcinoma. *International journal of oncology* 45(4):1547-1555.

45. Luo YD, Fang L, Yu HQ, Zhang J, Lin XT, Liu XY, Wu D, Li GX, Huang D, Zhang YJ et al (2021) p53 haploinsufficiency and increased mTOR signalling define a subset of aggressive hepatocellular carcinoma. *Journal of hepatology* 74(1):96-108.
46. Jin M, Wang J, Ji X, Cao H, Zhu J, Chen Y, Yang J, Zhao Z, Ren T, Xing J (2019) MCUR1 facilitates epithelial-mesenchymal transition and metastasis via the mitochondrial calcium dependent ROS/Nrf2/Notch pathway in hepatocellular carcinoma. *Journal of experimental & clinical cancer research* 38(1):136.
47. Ye YC, Zhao JL, Lu YT, Gao CC, Yang Y, Liang SQ, Lu YY, Wang L, Yue SQ, Dou KF et al (2019) NOTCH Signaling via WNT Regulates the Proliferation of Alternative, CCR2-Independent Tumor-Associated Macrophages in Hepatocellular Carcinoma. *Cancer research* 79(16):4160-4172.
48. Kitamura T, Qian BZ, Pollard JW (2015) Immune cell promotion of metastasis. *Nature reviews Immunology* 15(2):73-86.
49. Galdiero MR, Bonavita E, Barajon I, Garlanda C, Mantovani A, Jaillon S (2013) Tumor associated macrophages and neutrophils in cancer. *Immunobiology* 218(11):1402-1410.
50. Li X, Yao W, Yuan Y, Chen P, Li B, Li J, Chu R, Song H, Xie D, Jiang X et al (2017) Targeting of tumour-infiltrating macrophages via CCL2/CCR2 signalling as a therapeutic strategy against hepatocellular carcinoma. *Gut* 66(1):157-167.
51. Zongyi Y, Xiaowu L (2020) Immunotherapy for hepatocellular carcinoma. *Cancer letters* 470:8-17.

Figures

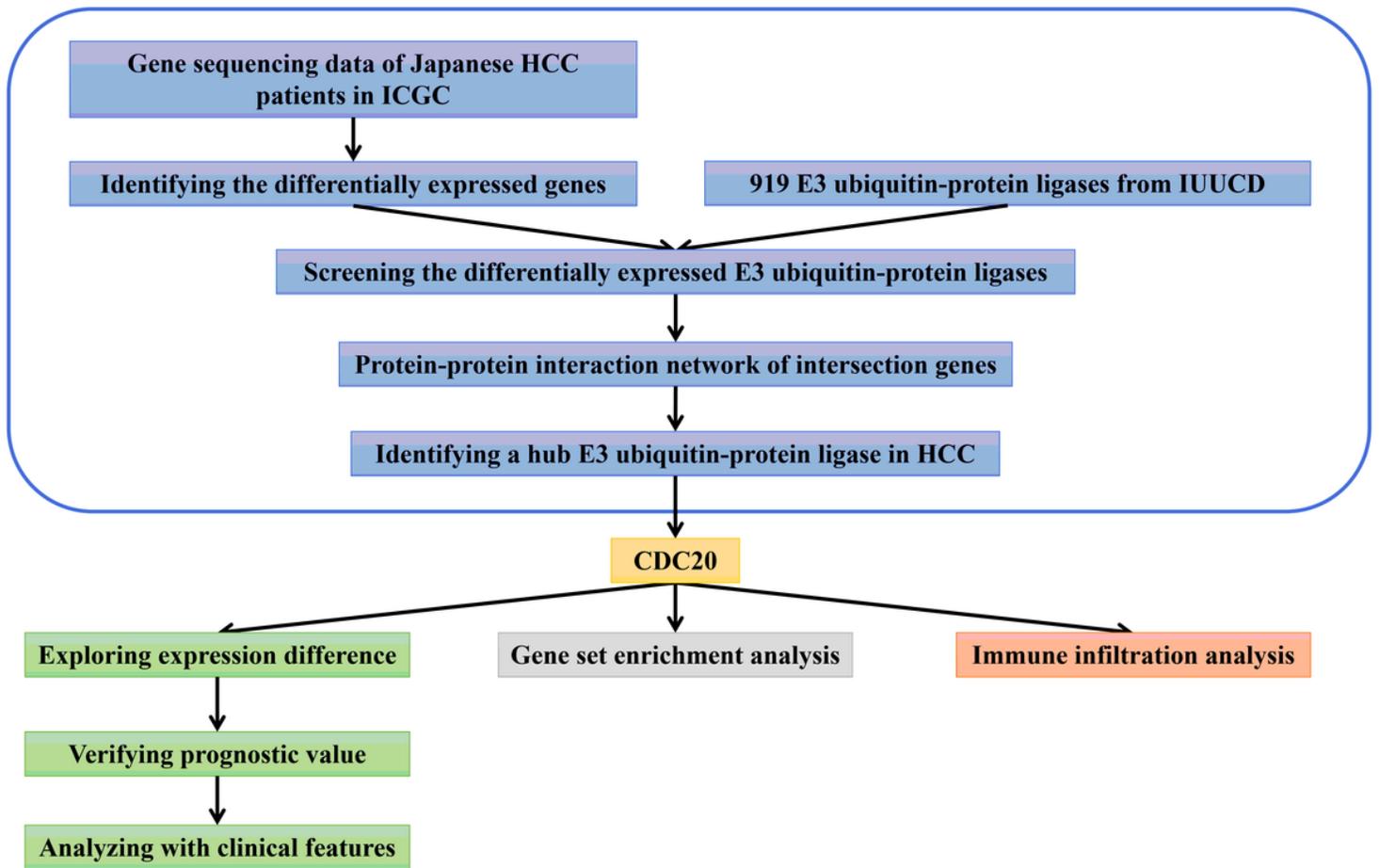


Figure 1

The processes of the entire research. The part inside the box was to identify CDC20 as a hub E3 ubiquitin-protein ligase in HCC.

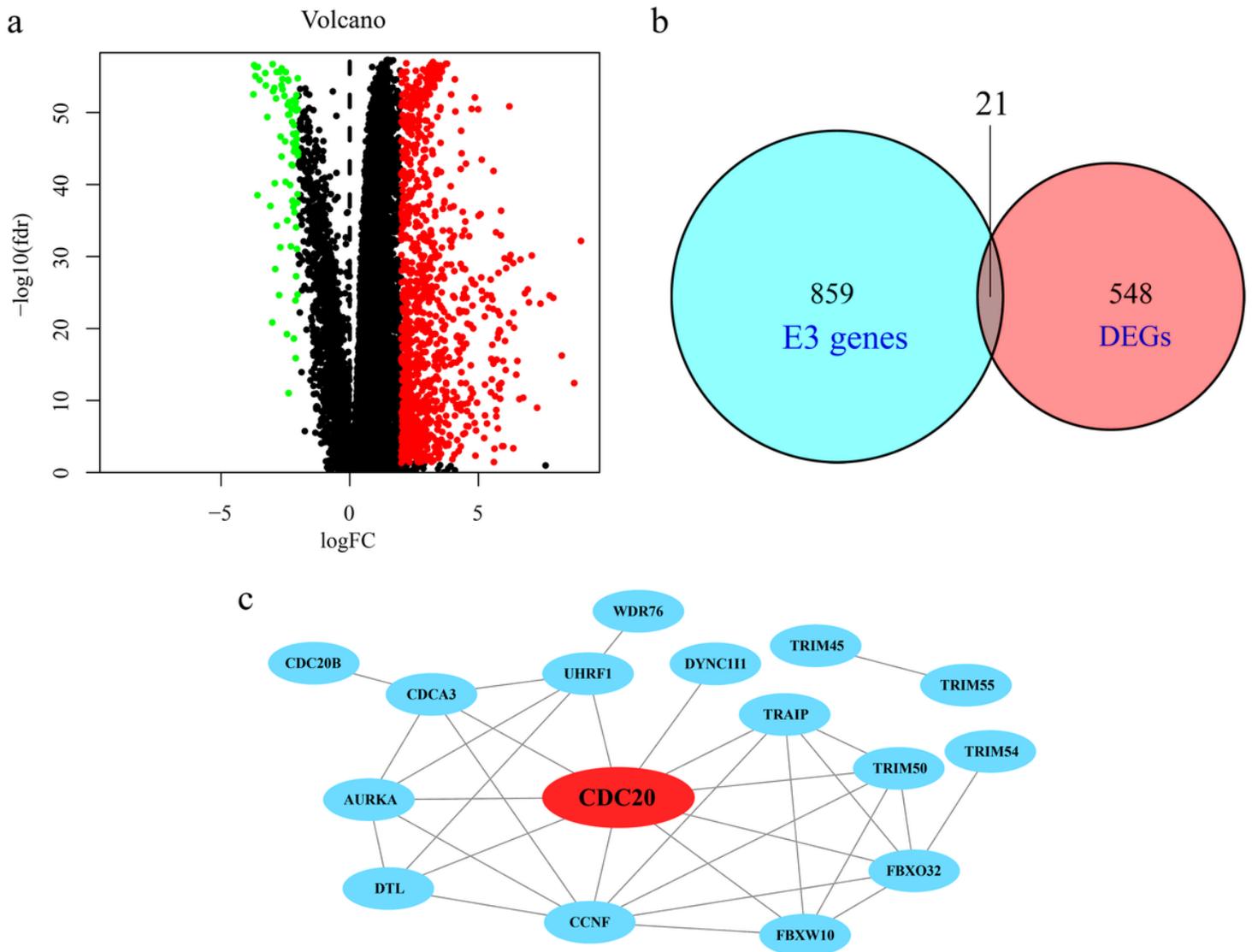


Figure 2

The procedures of identifying CDC20. a A volcano map including 859 DEGs in ICGC ($\log_{2}FC > 2$), b A Venn diagram to show the intersection of 859 DEGs in ICGC and 919 E3 ubiquitin-protein ligases, c PPI network of 16 intersection genes (degrees > 0) plotted by Cytoscape.

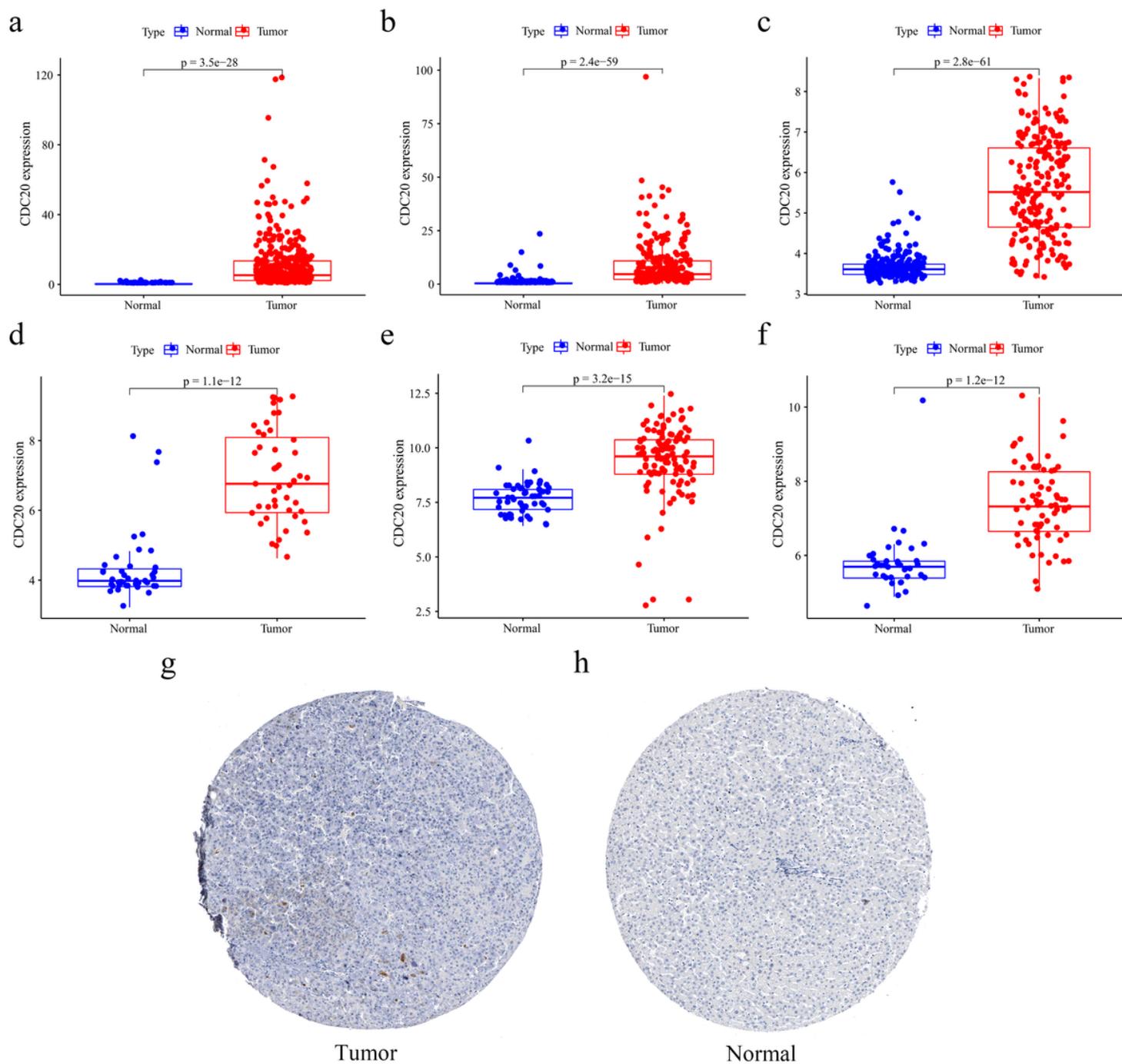


Figure 3

Differential expressions of CDC20 between HCC and normal tissues in different databases. a Differential expressions of CDC20 in TCGA, b Differential expressions of CDC20 in ICGC, c Differential expressions of CDC20 in GEO (GSE14520), d Differential expressions of CDC20 in GEO (GSE45267), e Differential expressions of CDC20 in GEO (GSE76427), f Differential expressions of CDC20 in GEO (GSE121248), g-h Immunohistochemical figures of CDC20 in HCC and normal tissue from HPA database.

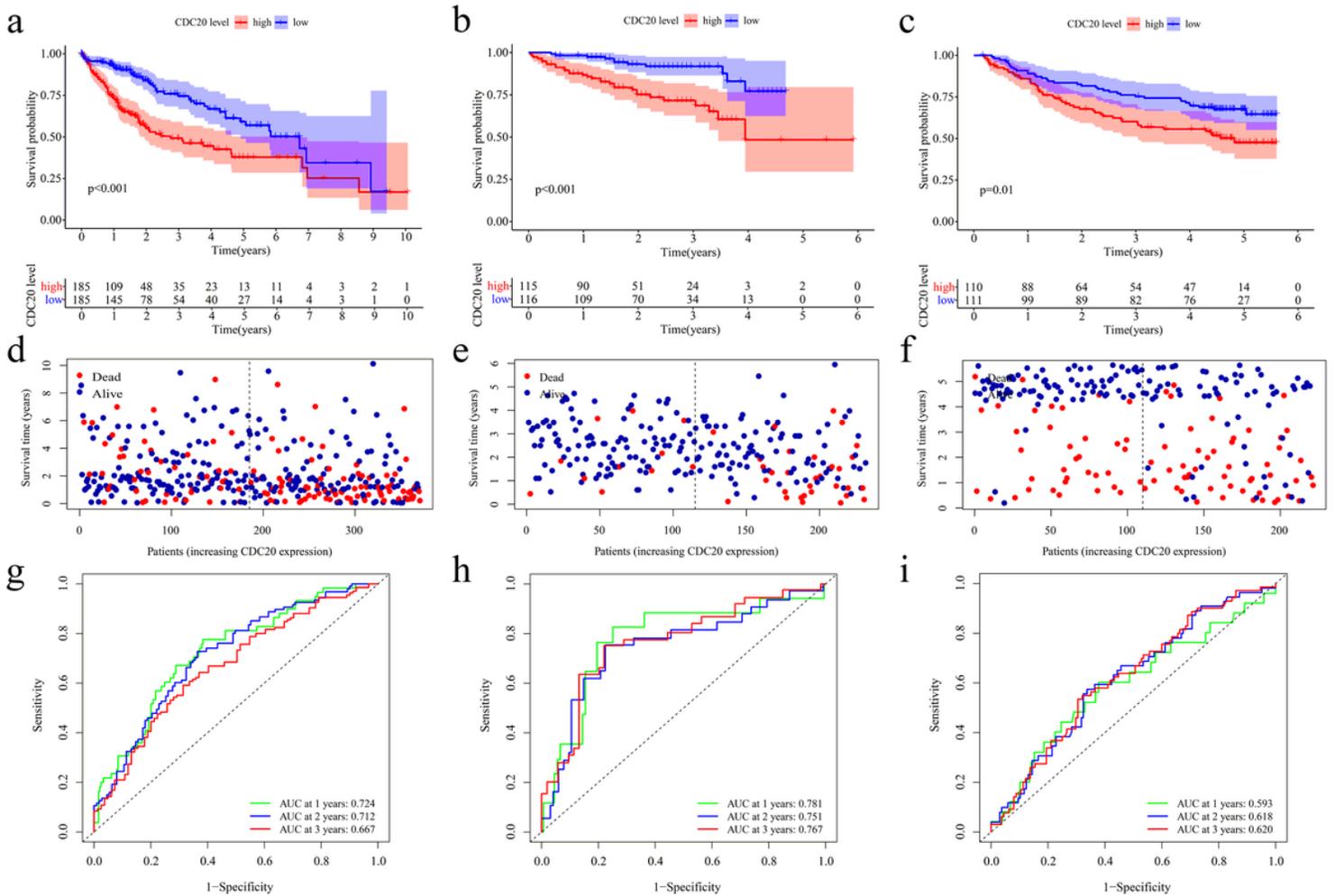


Figure 4

Illustrating the prognostic value of CDC20 in HCC. a-c The K-M curves of CDC20 by the data of HCC patients in TCGA, ICGC, and GEO (GSE14520) datasets, respectively, d-f The scatterplots to exhibit HCC patients' survival time and survival condition from TCGA, ICGC, and GEO (GSE14520) datasets, respectively, g-i The ROC curves that exhibited the AUC value of CDC20 at 1 years, 2 years, and 3 years based on data from TCGA, ICGC, and GEO (GSE14520) datasets, respectively.

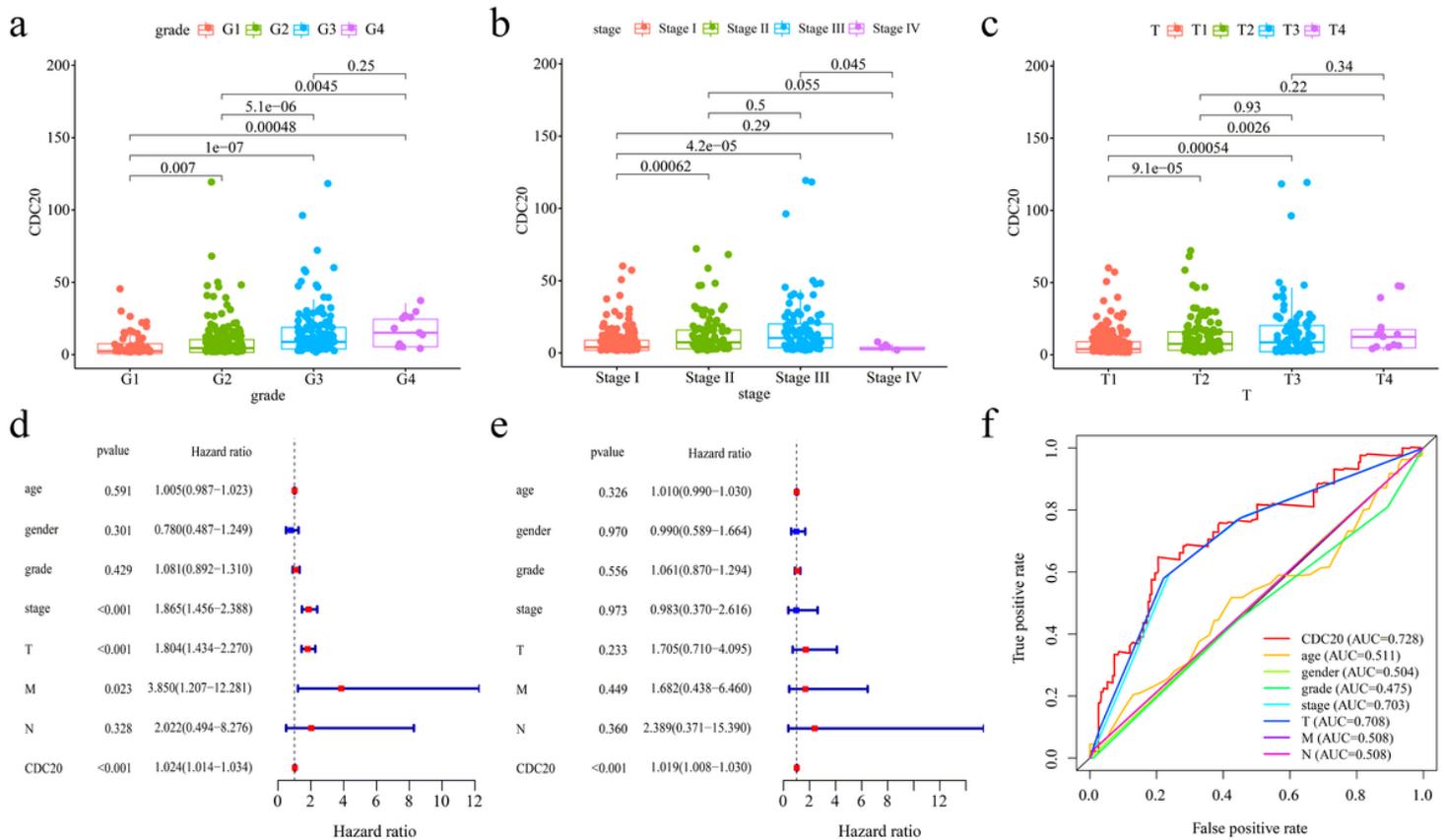


Figure 5

Analyzing CDC20 with other clinical features. a-c Correlations between the expression of CDC20 and clinical features including pathological grade, TNM stage, and T stage via the data of HCC patients from TCGA, d-e Univariate Cox and Multivariate Cox regression analyses between the expression of CDC20 and other clinical features, f ROC curves that compared the prognosis predictive value of the CDC20 and other clinical features in HCC.

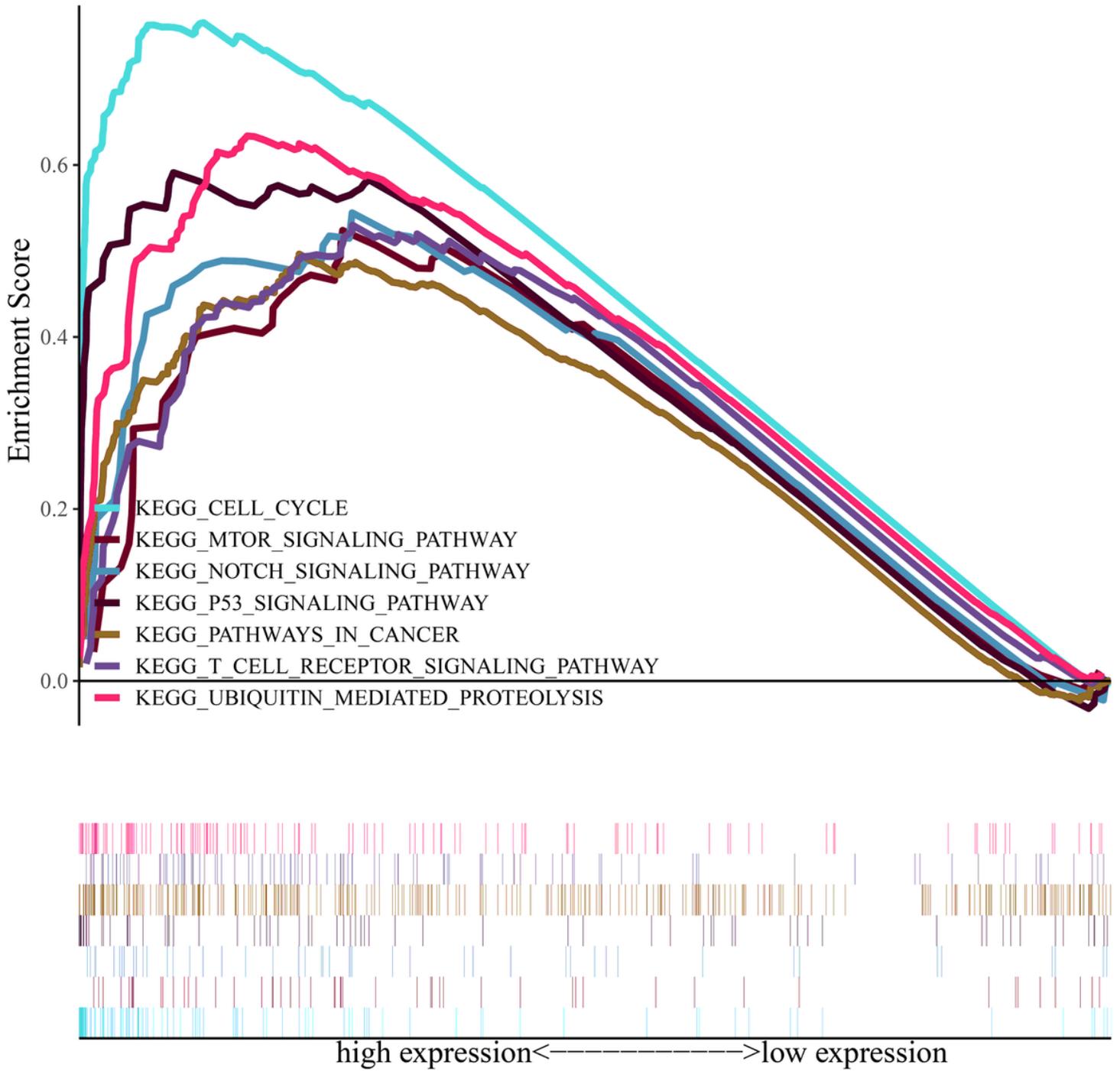


Figure 6

Gene set enrichment analysis of CDC20 in HCC.

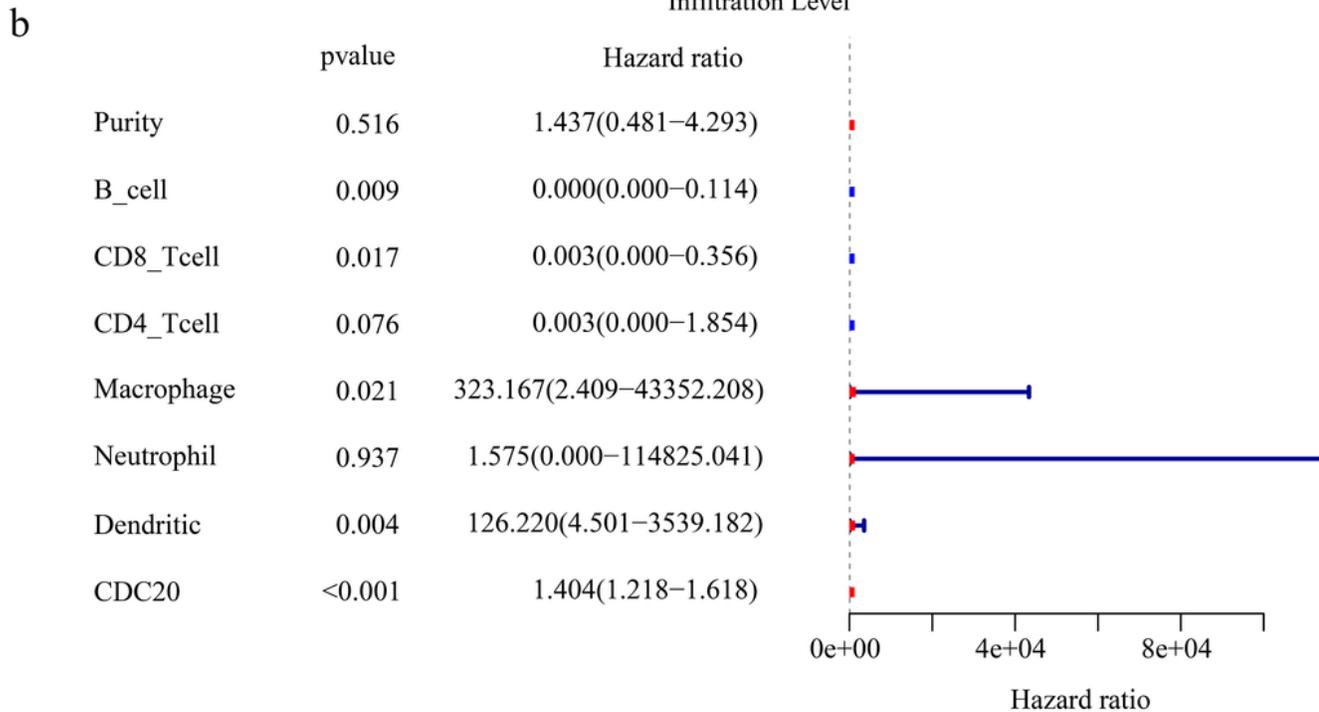
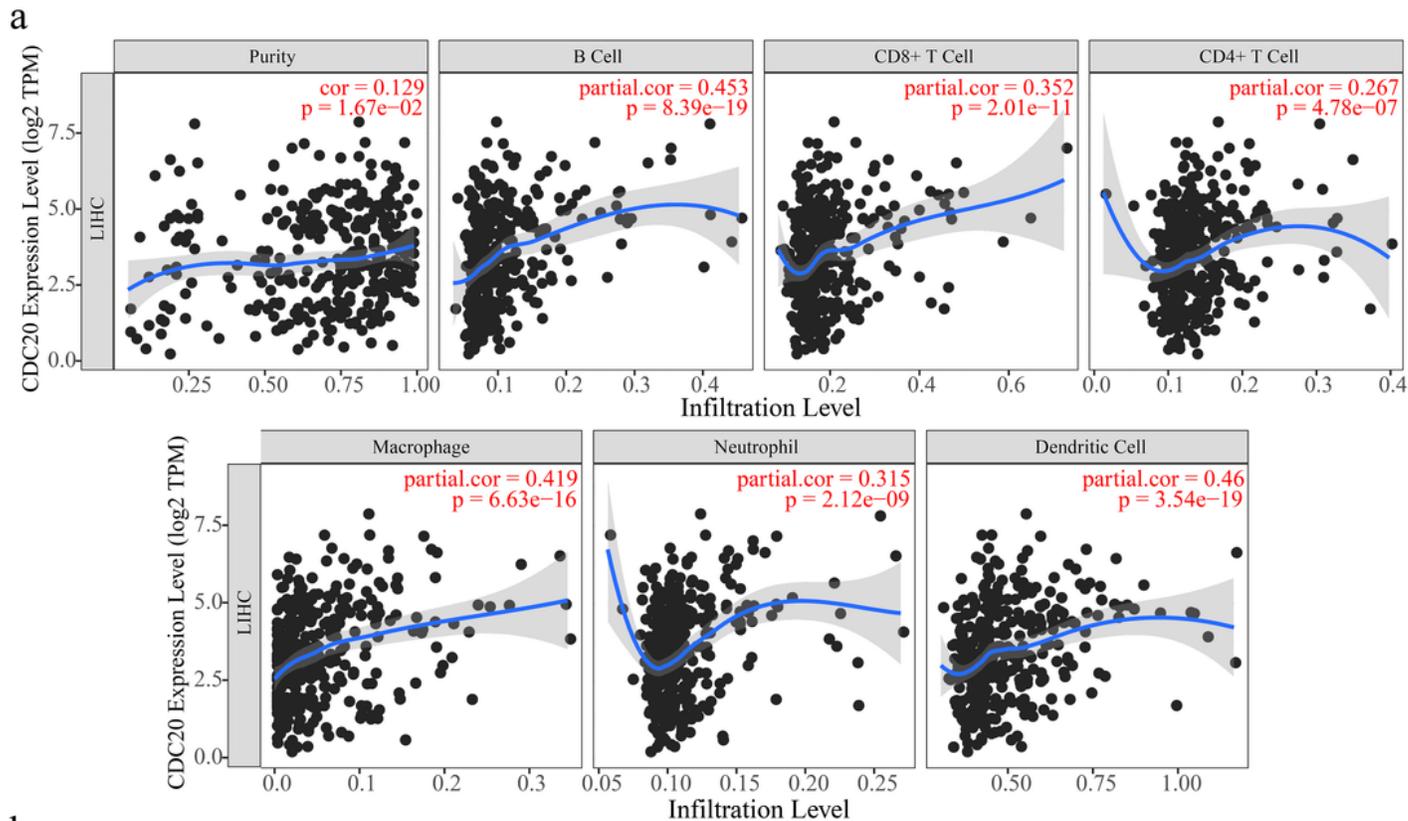


Figure 7

Analyzing CDC20 with 6 kinds of immune cells in HCC via TIMER. a Immune infiltrates analysis between CDC20 and 6 kinds of immune cells, b Multivariate Cox regression analyses between the expression of CDC20 and 6 kinds of immune cells.

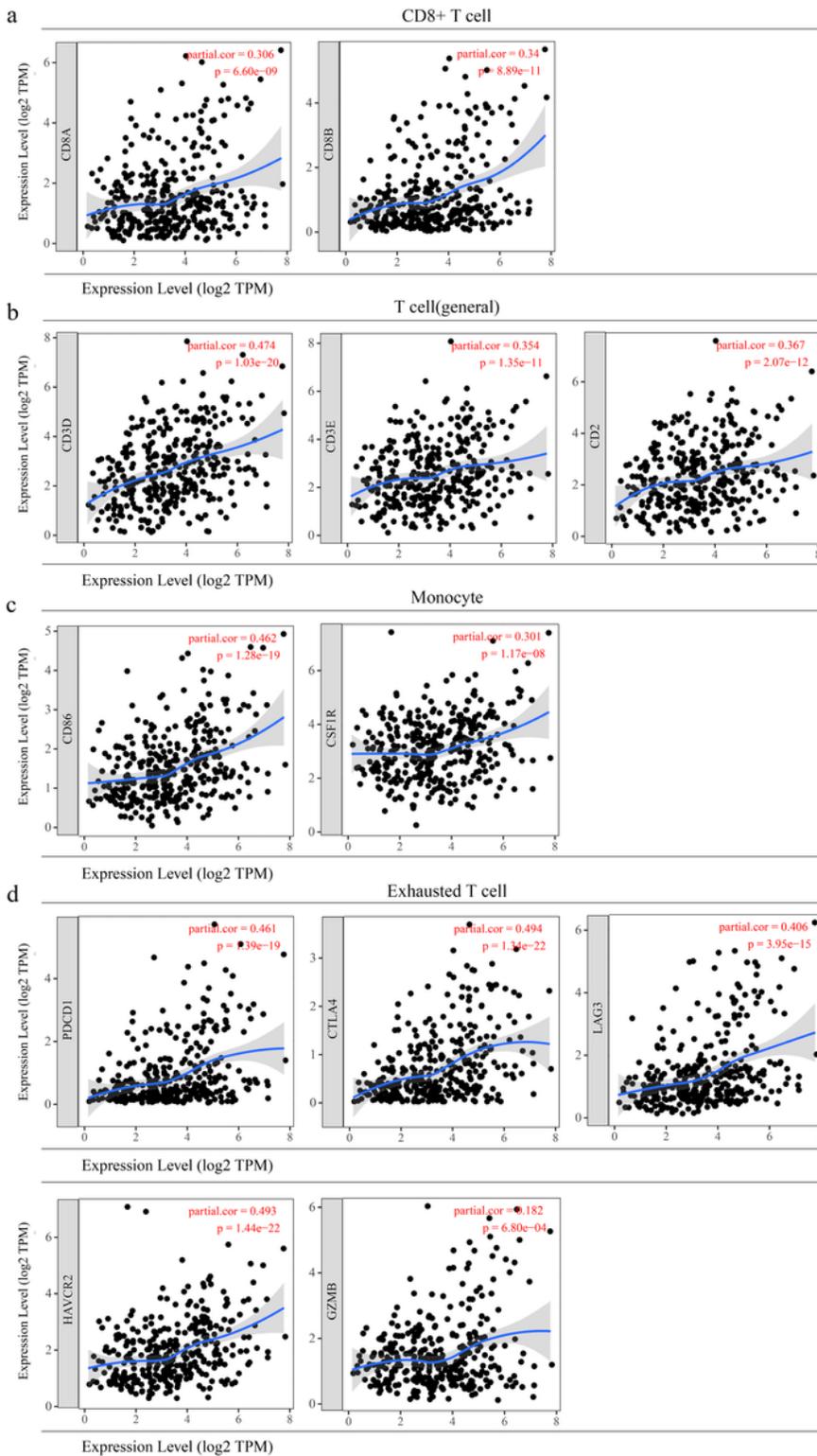


Figure 8

Analyzing the associations between CDC20 and the gene biomarkers of tumor-infiltrating immune cells in HCC via TIMER. a The scatter plots of the associations between CDC20 and the gene biomarkers of CD8+ T cell (CD8A and CD8B), b The scatter plots of the associations between CDC20 and the gene biomarkers of T cell (general) (CD3D, CD3E, and CD2), c The scatter plots of the associations between CDC20 and

the gene biomarkers of Monocyte (CD86 and CSF1R), d The scatter plots of the associations between CDC20 and the gene biomarkers of Exhausted T cell (PD-1, CTLA4, LAG3, TIM-3, and GZMB).