

Highly expressed centromere protein L indicates adverse survival and associates with immune infiltration in hepatocellular carcinoma

Zhili Zeng^{1,*}, Xiao Jiang^{1,*}, Zhibin Pan^{6,*}, Ruisheng Zhou^{1,*}, Zhuangteng Lin^{3,*}, Ying Tang^{2,4}, Ying Cui⁷, Enxin Zhang^{4,5}, Zebiao Cao¹

¹The First School of Clinical Medicine, Guangzhou University of Chinese Medicine, Guangzhou, Guangdong 510405, PR China

²Department of Oncology, Lingnan Medical Research Center of Guangzhou University of Chinese Medicine, Guangzhou, Guangdong 510405, PR China

³Department of Medical Technologic, The First Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, Guangdong 518000, PR China

⁴Department of Oncology, The First Affiliated Hospital of Guangzhou University of Chinese Medicine, Guangzhou, Guangdong 518000, PR China

⁵Department of Oncology, Shenzhen Hospital of Guangzhou University of Chinese Medicine, Shenzhen 518000, Guangdong, PR China

⁶Foshan Hospital of Traditional Chinese Medicine, Guangzhou University of Chinese Medicine, Foshan 528000, Guangdong, PR China

⁷Department of Psychiatry, The Third Affiliated Hospital of Guangzhou Medical University, Guangzhou, Guangdong 510150, PR China

*Equal contribution

Correspondence to: Ying Tang, Ying Cui, Enxin Zhang, Zebiao Cao; **email:** tangying@gzucm.edu.cn; 409765764@qq.com, <https://orcid.org/0000-0001-8952-9944>; ergep53@126.com, <https://orcid.org/0000-0001-6597-2001>; caozb3882@163.com, <https://orcid.org/0000-0003-4650-803X>

Keywords: centromere protein L, hepatocellular carcinoma, tumor immunity, prognosis, biomarker

Received: May 10, 2021

Accepted: September 11, 2021

Published: October 4, 2021

Copyright: © 2021 Zeng et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/3.0/) (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

ABSTRACT

Background: Hepatocellular carcinoma (HCC) is characterized by rapid progression, high recurrence rate and poor prognosis. Early prediction for the prognosis and immunotherapy efficacy is of great significance to improve the survival of HCC patients. However, there is still no reliable predictor at present. This study is aimed to explore the role of centromere protein L (CENPL) in predicting prognosis and its association with immune infiltration in HCC.

Methods: The expression of CENPL was identified through analyzing the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) data. The association between CENPL expression and clinicopathological features was investigated by the Wilcoxon signed-rank test or Kruskal Wallis test and logistic regression. The role of CENPL in prognosis was examined via Kaplan-Meier method and Log-rank test as well as univariate and multivariate Cox regression analysis. Besides, in TIMER and GEPIA database, we investigated the correlation between CENPL level and immunocyte and immunocyte markers, and the prognostic-related methylation sites in CENPL were identified by MethSurv.

Results: CENPL had a high expression in HCC samples. Increased CENPL was prominently associated with unfavorable survival, and maybe an independent prognostic factor of worse overall survival (OS), disease-specific survival (DSS), disease-free interval (DFI), progression-free interval (PFI). Additionally, CENPL expression was significantly correlated with immune cell infiltration and some markers. CENPL also contained a methylation site that was notably related to poor prognosis.

Conclusions: Elevated CENPL may be a promising prognostic marker and associate with immune infiltration in HCC.

INTRODUCTION

Primary liver cancer is a rapidly developing and aggressive malignant tumor with high incidence rate and short survival time. According to the statistics of the International Agency for Research on Cancer on global cancer incidence and mortality in 2020, primary liver cancer is the sixth most commonly diagnosed cancer and the third leading cause of cancer death worldwide in 2020, with approximately 906,000 new cases and 830,000 deaths and it ranks the second in the mortality rate of male cancer [1]. Hepatocellular carcinoma (HCC) is the main type of primary liver cancer [2]. Although early liver cancer can be cured by surgery or liver transplantation, more than 80% of patients have no chance of surgery. Even with successful surgical excision or liver transplantation, the postoperative recurrence is still common and elusive. Especially, early recurrence and metastasis are often difficult to be detected, so the 5-year survival rate after surgical resection and liver transplantation are only 36–70% and 60–70%, respectively [3]. The majority of patients are found in the late stage or after metastasis, while the treatment in the late stage is limited and the therapeutic effect is poor. With the clinical application of immune checkpoint inhibitors (ICIs), new dawn has been brought to tumor patients, but only a few patients have experienced clinical benefits [4]. Therefore, early prediction for the occurrence and progress as well as the immunotherapy efficacy is of great significance to improve the prognosis of HCC patients and avoid unnecessary drug toxicity. However, to date, there has been no robust predictor for HCC.

Centromere protein L (CENPL) is a member of the centromere protein (CENP) family, which is necessary for normal cell division (mitosis and meiosis) [5]. We all know that centromeres are specialized DNA sequences that connect a pair of sister chromatids, and are mainly regarded as loci that guide chromosomal behavior. The kinetochore is a cell structure attached to the centromere. During mitosis, the two sister chromatids are drawn to the poles of the cell through the spindle filaments (stellar rays of the spindle) attached to the kinetochore. If the centromere or kinetochore is abnormal, the chromosomes will be randomly allocated into the daughter cells during cell division, and the chromosome number will change. The centromere is composed of more than 15 centromere-specific proteins binding together, including the CENP family [5]. In addition, the CENP family is also the basis of

kinetochore assembly and function, and they determine the correct separation of chromosomes [5, 6]. Otherwise, it will lead to abnormal chromosome number and induce tumorigenesis. Studies have shown that most cancer cells are aneuploidy [7, 8]. CENPL can be combined with CENPN to identify CENPA nucleosomes, and it plays a vital role in the process of recruiting other centromere components to assemble into centromeres [9, 10]. Previous researches have reported the role of CENPA [11, 12], CENPE [13], CENPF [14] in HCC, however, there is no research on the role of CENPL in HCC and other cancers. Based on the Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO) databases, we analyzed the expression of CENPL in HCC, and explored the possible pathways by which abnormally expressed CENPL was involved in the occurrence and progression of HCC. Moreover, we also explored the relationship between CENPL expression and immune cell infiltration, and investigated the abnormal methylation sites of CENPL.

MATERIALS AND METHODS

Gene expression profiles and clinical information

Using TCGA database (<https://portal.gdc.cancer.gov/repository>), we obtained the gene expression data (424 cases, Workflow Type: HTSeq-FPKM) and corresponding clinical information. Samples with deficient or unclear information on important clinicopathological characteristics were excluded. In all, 374 HCC samples and 50 adjacent normal samples were brought into our research. Patient's clinical information contained age, gender, family history, clinical stage, TNM stage, AFP, new tumor event, residual tumor, vascular invasion, postoperative ablation embolization, tumor status, histologic grade, Child-Pugh and risk factor (viral hepatitis and/or alcohol consumption). In addition, we used gene expression profiles of GSE121248 and GSE54236, which were downloaded from the GEO database, to confirm the CENPL expression in HCC. Moreover, we validated the protein level of CENPL through obtaining the corresponding immunohistochemical (IHC) images in Human Protein Atlas (HPA) database (<http://www.proteinatlas.org/>). HPA is an open access knowledge repository that integrates the work of many laboratories and technology platforms around the world to explore and annotate more than 20,000 human genes in detail at the protein level [15, 16]. Image-Pro Plus software (version 6.0; Media

Cybernetics, Inc.) was applied to detect the mean integrated optical density (IOD) value of IHC images. The IOD value, representing the protein level of CENPL, was statistically analyzed by non-paired *T*-test in the GraphPad Prism® version 8.0 software. $P < 0.05$ indicates statistical significance.

Gene Set Enrichment Analysis (GSEA)

A computational method called Gene Set Enrichment Analysis (GSEA) can be used to decide whether an *a priori* defined set of genes has statistical significance, concordant differences between two biological states [17]. The level of the gene was recognized as a phenotype label. The number of gene set permutations was 1,000 times for each analysis. Biological pathways with $P < 0.05$ and false discovery rate (FDR) < 0.05 were considered to be significantly enriched.

Kaplan-Meier plotter analysis

The Kaplan-Meier plotter [18] (<http://www.kmplot.com/analysis/index.php?p=background>) was applied to analyze the correlation between CENPL expression and survival outcome in lung cancer, ovarian cancer, breast cancer and gastric cancer. Log-rank *P*-values < 0.05 represents statistical significance.

Co-expression, GO and KEGG enrichment analyses

Coexpedia (<https://www.coexpedia.org/index.php>) is a unique network tool based on the human and mouse gene chip data in the GEO database. It can analyze the functional, biological, and medical correlations between genes through statistical analysis, and then construct a co-expression network [19]. It was used to find the co-expressed genes of CENPL. Further, the co-expressed genes were utilized for Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis in the Database for Annotation, Visualization and Integrated Discovery (DAVID) database (version 6.8; <https://david.ncifcrf.gov/>) [20]. The annotation and functional analyses of GO consist of biological process (BP), cellular component (CC) and molecular function (MF). False discovery rate (FDR) < 0.05 was set as the enrichment cut-off value.

Immunocyte infiltration

Tumor Immune Estimation Resource (TIMER) is an online network platform for comprehensive analysis of the immunocyte infiltration abundance (<http://timer.cistrome.org/>). It can predict the immunocyte infiltration levels in over 10000 tumors from 32 tumor types, and analyze the associations between the immunocyte infiltration abundance and

gene expression, survival time, and other clinicopathological factors [21]. We searched the expression of CENPL in various cancers and uncovered the correlation between CENPL expression and six main immune cell types as well as their markers in TIMER. Furthermore, we also applied TIMER's multivariate Cox analysis to explore the effect of immune cell infiltration and CENPL expression on survival outcome.

GEIPA analysis

Gene Expression Profiling Interactive Analysis (GEPIA, <http://gepia.cancer-pku.cn/index.html>) is a network tool based on TCGA and GTEx databases. It provides customizable functions such as tumor/normal differential expression analysis, profiling according to cancer types or pathological stages, patient survival analysis, similar gene detection, correlation analysis and dimensionality reduction analysis [22]. GEPIA was employed to validate the transcriptional level of CENPL in HCC and other cancers, and to demonstrate the relationship between CENPL expression and the major markers of immune cells.

DNA methylation

MethSurv (<https://github.com/vijayachitrabio/MethSurv>) is a web tool for univariate and multivariate survival analysis based on DNA methylation data from 25 different cancer types and 7,358 patients in the TCGA database [23]. Using the MethSurv, CENPL DNA methylation sites related to the survival was disclosed.

Statistical analysis

The expression differences of CENPL in HCC samples and normal samples, HCC samples and adjacent samples were analyzed by Wilcoxon rank-sum tests and Wilcoxon signed-rank tests, respectively. The Wilcoxon signed-rank test or Kruskal Wallis test and logistic regression were used to investigate the relationship between CENPL expression and clinicopathological factors. The role of CENPL in survival was examined via Kaplan-Meier method and Log-rank test as well as univariate and multivariate Cox regression analysis. All statistical analyses were conducted on R (version 3.6.1, 2019-07-05, R Foundation, Vienna, Austria), $P < 0.05$ was set as the statistical threshold. The median expression value of CENPL was regarded as the cut-off value.

Ethics approval and consent to participate

No ethics approval was required for this work. All utilized public data sets were generated by others who obtained ethical approval.

Availability of data and materials

The datasets generated and/or analyzed during the current study are available in the GSE54236 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE54236>), and GSE121248 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE121248>) from the Gene Expression Omnibus (GEO) database (<https://www.ncbi.nlm.nih.gov/geo/>).

RESULTS

Clinicopathological features of patients

As showed in Table 1, the clinical information of 374 HCC patients was extracted from the TCGA database in August 2020. After excluding the samples with missing or unclear clinicopathological features, finally, there were 121 females and 250 males. Only 9% ($n = 34$) patients were under 40 years old and 35.0% ($n = 112$) patients had HCC family history. Besides, most patients (65.8%, $n = 231$) had risk factors, such as alcohol consumption and/or viral hepatitis. The histologic grade included 232 (63.4%) cases in G1-2 and 134 (36.6%) cases in G3-4. The clinical stage was composed of 74.1% ($n = 257$) cases in stage I-II and 25.9% ($n = 90$) cases in stage III-IV. There were 275 of 368 (74.7%) patients in T1-2 and 93 of 368 (25.3%) patients in T3-4. Only 1.6% (4 of 256), 1.5% (4 of 270), and 34.6% (109 of 315) cases were found to have lymph node metastasis, distant metastasis, and vascular invasion, respectively. The distribution of Child-Pugh involved 90.8% Child-Pugh A patients ($n = 217$) and 9.2% Child-Pugh B-C patients ($n = 22$). 52.9% ($n = 147$) patients were in $\text{AFP} < 20$, 23.7% ($n = 66$) in $20 \leq \text{AFP} < 400$, and 23.4% ($n = 65$) in $\text{AFP} \geq 400$. 97.7% (338 of 346) and 91.9% (319 of 347) patients did not undergo radiation therapy and postoperative chemotherapy, respectively. After operation, there were 5.3% ($n = 18$) patients with residual tumor and 48.3% ($n = 169$) patients with new tumor event. Until we extracted the data, a total of 57.1% ($n = 201$) patients were in a tumor-free state.

The expression of CENPL in HCC and other cancers

In our study, CENPL expression was compared between 374 HCC samples and 50 adjacent normal samples. We found that the expression of CENPL was significantly elevated in HCC (Figure 1A, 1B). This result was confirmed in GSE121248 and GSE54236 datasets (Figure 1C, 1D). Further, the TIMER and GEPIA analysis showed that the transcriptional levels of CENPL in various cancer types, such as HCC, lung cancer, pancreatic cancer, ovarian cancer, breast cancer, gallbladder cancer, cervical cancer, esophageal cancer,

glioblastoma multiforme Tumor, bladder urothelial carcinoma, were significantly higher than that in normal tissues (Supplementary Figures 1 and 2). Moreover, we found that the protein level of CENPL was also significantly increased in HCC based on the HPA database (Figure 1E–1G). These results suggested that high mRNA and protein levels of CENPL in HCC were consistent in different databases.

Association between CENPL and clinicopathological factors

As we can see in Figure 2A–2I, up-regulated CENPL was distributed in patients with family history (Yes vs. No, $p = 0.011$), vascular invasion (No vs. Yes, $p = 0.032$), new tumor event ($p = 0.033$), risk factor (R1 vs. R2 vs. R3 vs. R4, $p = 0.043$), with tumor ($p = 0.006$), high AFP ($\text{AFP} < 20$ vs. $20 \leq \text{AFP} < 400$ vs. $\text{AFP} \geq 400$, $p = 2.655e-08$) as well as unfavorable histologic grade (G1-2 vs. G3-4, $p = 4.234e-10$), clinical stage (Stage I–II vs. Stage III–IV, $p = 0.008$) and topography (T) stage (T1-2 vs. T3-4, $p = 0.005$). Similar but more specific results were obtained by logistic regression analysis (Table 2).

The prognostic value of enhanced CENPL

Figure 3 unveiled that enhanced CENPL was significantly related to poor overall survival (OS) ($p = 3.995e-05$), disease-specific survival (DSS) ($p = 3.423e-05$), disease-free interval (DFI) ($p = 0.012$), and progression-free interval (PFI) ($p = 7.493e-04$). Moreover, this similar phenomenon occurred in other cancers, such as OS, post progression survival (PPS), first progression (FP) in lung cancer; OS, PPS, progression-free survival (PFS) in ovarian cancer; OS, relapse-free survival (RFS), distance metastasis free survival (DMFS) in breast cancer (Supplementary Figure 3). Furthermore, univariate and multivariate Cox regression analysis found that elevated CENPL can independently predict worse OS (HR = 1.7, 95% CI [1.2–2.6], $p = 0.008$), DSS (HR = 2.2, 95% CI [1.4–3.6], $p = 0.001$), DFI (HR = 1.6, 95% CI [1.1–2.3], $p = 0.006$), and PFI (HR = 1.7, 95% CI [1.2–2.3], $p = 0.002$) in HCC (Table 3, Figure 4).

Pathways related to CENPL in HCC

GSEA unraveled that spliceosome, nucleotide excision repair, DNA replication, cell cycle, homologous recombination, ubiquitin mediated proteolysis, mismatch repair, p53 signaling pathway, oocyte meiosis and pyrimidine metabolism were significantly enriched in the high CENPL expression phenotype (Table 4; Figure 5A), indicating that elevated CENPL might

Table 1. HCC patient characteristics based on TCGA.

Clinical characteristics		Total	%
Age at diagnosis (years)	>40	336	91.0
	≤40	34	9.0
Gender	male	250	67.6
	female	121	32.4
Family history	Yes	112	35.0
	No	208	65.0
Histologic grade	G1-G2	232	63.4
	G3-G4	134	36.6
Clinical stage	I-II	257	74.1
	III-IV	90	25.9
T	T1-T2	275	74.7
	T3-T4	93	25.3
N	N0	252	98.4
	N1	4	1.6
M	M0	266	98.5
	M1	4	1.5
Residual tumor	R0	324	94.7
	R1	18	5.3
Tumor status	tumor free	201	57.1
	with tumor	151	42.9
Vascular invasion	Yes	109	34.6
	No	206	65.4
Child-Pugh	A	217	90.8
	B-C	22	9.2
AFP	AFP <20	147	52.9
	20 < AFP <400	66	23.7
	AFP ≥400	65	23.4
New tumor event	Yes	169	48.3
	No	181	51.7
Risk factor	Alcohol consumption and viral hepatitis	39	11.1
	Alcohol consumption	78	22.2
	Viral hepatitis	114	32.5
	Neither	120	34.2
Postoperative ablation embolization	Yes	28	8.1
	No	319	91.9
Radiation therapy	Yes	8	2.3
	No	338	97.7

Abbreviations: T: topography distribution; N: lymph node metastasis; M: distant metastasis; AFP: alpha fetal protein.

participate in the occurrence and progression of HCC through these pathways.

Additionally, we obtained the co-expressed genes of CENPL in Coexpedia database (Figure 6A). Among these genes, Nuf2 had the strongest positive correlation with CENPL in HCC and many other

common cancers (Supplementary Figure 4). Since co-expressed genes often have similar functions [24], we analyzed the enrichment pathways of these co-expressed genes via DAVID. For BP, co-expression genes were significantly enriched in kinetochore organization, mitotic sister chromatid segregation, spindle organization, kinetochore assembly, sister

chromatid cohesion DNA replication. For CC, co-expression genes were prominently enriched in condensed chromosome kinetochore, kinetochore, spindle microtubule, mitotic spindle, chromatin. For MF, co-expression genes were significantly enriched

in kinetochore binding, microtubule motor activity, damaged DNA binding, microtubule binding, chromatin binding, ATP binding and DNA binding (Figure 6B). KEGG pathway enrichment analysis showed that co-expression genes were

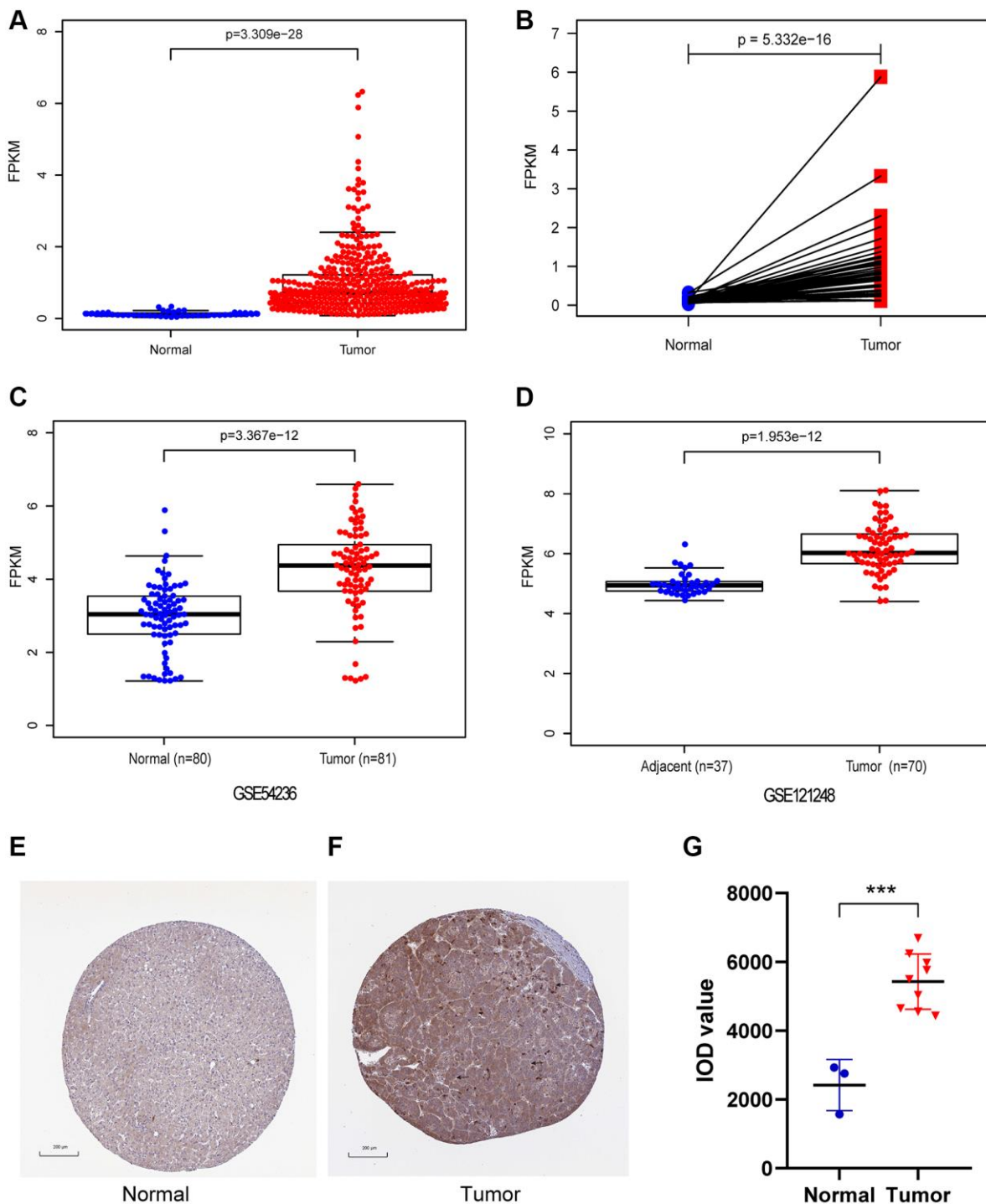


Figure 1. The mRNA and protein expression of CENPL in HCC. (A) CENPL showed significantly high expression in HCC tissues than in normal tissues using Wilcoxon rank sum test. (B) The expression of CENPL was prominently highly expressed in HCC tissues compared with adjoining non-cancerous tissues via Wilcoxon signed-rank test. (C and D) suggested CENPL was prominently enhanced in HCC samples from GSE121248 and GSE54236. (E–G) compared with the normal group, the protein level of CENPL in the liver cancer group was significantly increased. Abbreviation: CENP: centromere protein.

significantly enriched in cell cycle, DNA replication, oocyte meiosis, pyrimidine metabolism, mismatch repair, p53 signaling pathway, purine metabolism, nucleotide excision repair and base excision repair (Figure 6C).

The effect of CENPL expression on immune cell infiltration

Based on the TIMER and GEPIA database, we found that CENPL expression was positively correlated with

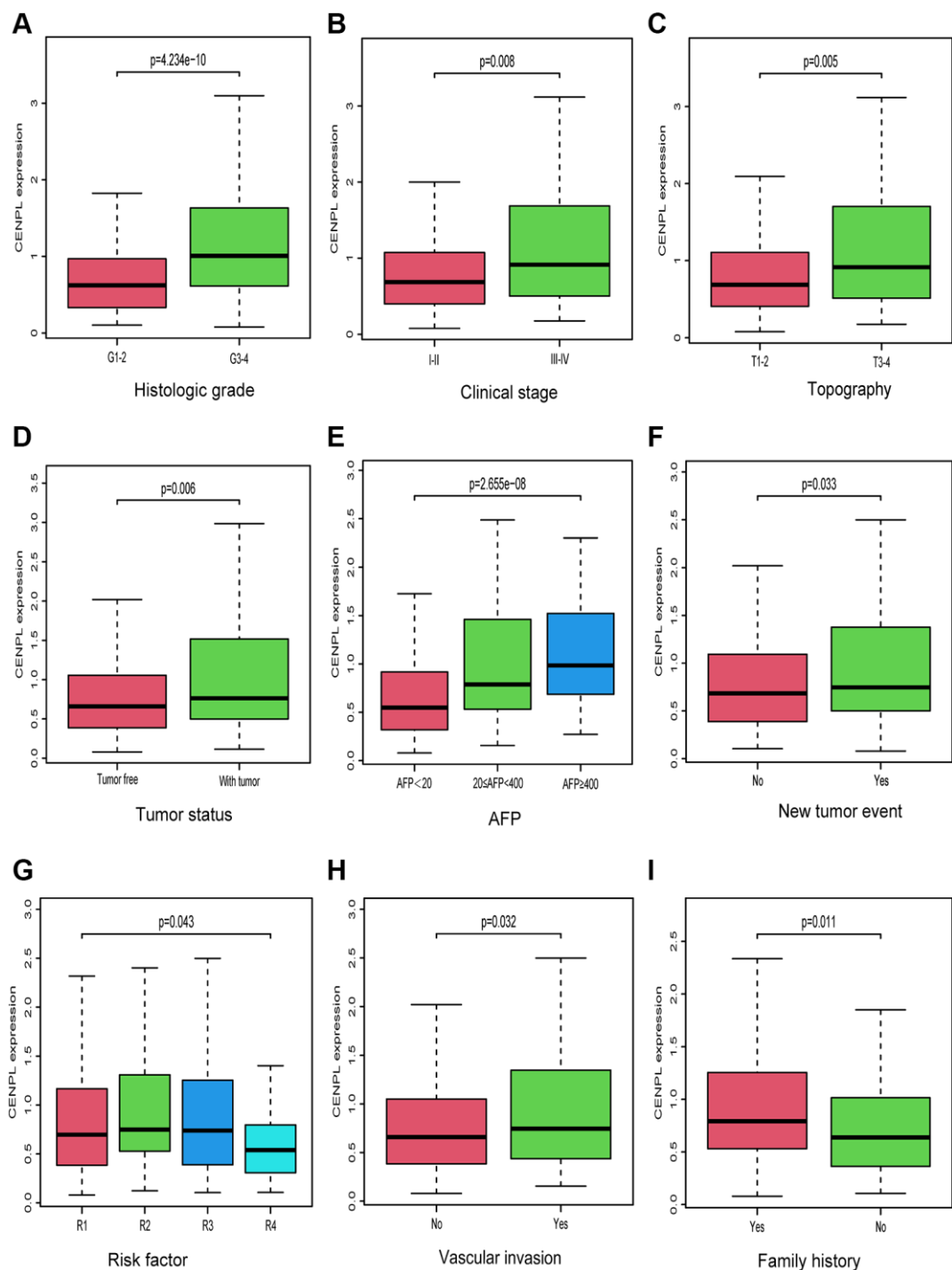


Figure 2. Relationship between CENPL transcription level and clinicopathologic characteristics. Increased CENPL had significant correlation with (A) histologic grade, (B) clinical stage, (C) topography, (D) tumor status, (E) AFP, (F) New tumor event, (G) Risk factor, (H) Vascular invasion, (I) Family history. Abbreviations: CENP: centromere protein; T: topography distribution; N: lymph node metastasis; M: distant metastasis; AFP: alpha fetal protein.

Table 2. Association between CENPL expression and clinicopathologic characteristics (logistic regression).

Clinical characteristics	Total	OR	P-value
Age at diagnosis (years) (>40 vs. ≤40)	370	0.9 (0.4–1.8)	0.719
Gender (female vs. male)	371	0.6 (0.4–1.0)	0.033
Family history (no vs. yes)	320	0.6 (0.4–0.9)	0.020
Histologic grade (G3-4 vs. G1-2)	366	3.4 (2.2–5.4)	0.000
Clinical stage (III-IV vs. I-II)	347	1.7 (1.1–2.8)	0.026
T (T3-4 vs. T1-2)	368	1.8 (1.1–3.0)	0.012
N (N1 vs. N0)	256	1.0 (0.1–8.4)	1.000
M (M1 vs. M0)	270	1.0 (0.1–8.4)	1.000
Residual tumor (R1-2 vs. R0)	342	1.6 (0.6–4.5)	0.337
Tumor status (with tumor vs. tumor free)	352	1.6 (1.0–2.4)	0.041
Vascular invasion (yes vs. no)	315	1.5 (1.0–2.5)	0.070
Child-Pugh (B-C vs. A)	239	1.6 (0.6–4.5)	0.337
New tumor event (yes vs. no)	350	1.4 (0.9–2.2)	0.109
Radiation therapy (yes vs. no)	346	0.6 (0.1–2.5)	0.479
Postoperative ablation embolization (yes vs. no)	347	1.2 (0.5–2.6)	0.682
Risk factor			
RF2 vs. RF1	235	1.3 (0.8–2.2)	0.319
RF3 vs. RF1	199	1.2 (0.7–2.1)	0.600
RF4 vs. RF1	160	0.5 (0.2–1.1)	0.095
RF3 vs. RF2	192	0.9 (0.5–1.6)	0.712
RF4 vs. RF2	153	0.4 (0.2–0.9)	0.020
RF4 vs. RF3	117	0.5 (0.2–1.0)	0.051
AFP			
AFP2 vs. AFP1	213	2.9 (1.6–5.4)	0.000
AFP3 vs. AFP1	131	4.0 (2.2–7.6)	0.000
AFP3 vs. AFP2	212	1.4 (0.7–2.9)	0.392

Abbreviations: OR: odd ratio; T: topography distribution; N: lymph node metastasis; M: distant metastasis; RF: risk factor (RF1: neither alcohol consumption or viral hepatitis; RF2: viral hepatitis; RF3: alcohol consumption; RF4: both alcohol consumption and viral hepatitis); AFP: alpha fetal protein (AFP1 represents AFP <20; AFP2 represents 20 ≤ AFP <400; AFP3 represents AFP ≥400).

the infiltration of B cell ($r = 0.409$, $p = 2.63e-15$), CD8+ T Cell ($r = 0.279$, $p = 1.64e-07$), CD4+ T Cell ($r = 0.394$, $p = 3.00e-14$), Macrophage ($r = 0.47$, $p = 4.01e-20$), Neutrophil ($r = 0.421$, $p = 3.04e-16$), Dendritic Cell ($r = 0.402$, $p = 1.32e-14$) (Figure 5B), as well as some immunocyte biomarkers (Supplementary Figure 5 and Table 5). In addition, TIMER multivariate Cox analysis exhibited that B cell (Coef = -7.514, HR = 0.00, $p = 0.045$), CD8 T cell (Coef = -5.056, HR = 0.01, $p = 0.046$), CD4 T cell (Coef = -9.241, HR = 0.00, $p = 0.016$) infiltration were negative independent predictors for poor prognosis, while macrophage (Coef = 8.546, HR = 5146.34, $p = 0.003$) and dendritic

(Coef = 4.945, HR = 140.42, $p = 0.012$) infiltration and CENPL expression were positive independent predictors for poor prognosis (Table 6), indicating low B cell, CD8 T cell, CD4 T cell infiltration and high macrophage and dendritic infiltration as well as CENPL expression predicted poor prognosis.

In TIMER database, CENPL expression is prominently positively correlated with markers of M1 Macrophage (IRF5: Cor = 0.478, $P = 3.86E-21$), M2 Macrophage (CD163: Cor = 0.194, $P = 2.91E-04$; VSIG4: Cor = 0.238, $P = 8.12E-06$; MS4A4A: Cor = 0.228, $P = 1.99E-05$), Treg (STAT5B: Cor = 0.379, $P = 3.29E-13$;

CCR8: $Cor = 0.504$, $P = 1.20E-23$), B cell and T cell, especially T-cell failure markers such as PD-1($Cor = 0.363$, $P = 3.60E-12$), CTLA4 ($Cor = 0.364$, $P = 3.02E-12$), LAG3 ($Cor = 0.278$, $P = 1.46E-07$), TOX ($Cor = 0.322$, $P = 9.49E-10$), TIGIT ($Cor = 0.367$, $P = 1.89E-12$), GZMB ($Cor = 0.137$, $P = 1.10E-02$) and Tim-3 ($Cor = 0.414$, $P = 9.94E-16$). Similar results were obtained in GEPIA database.

The methylation site of CENPL

We found two DNA methylation sites of CENPL, namely, CENPL-3'UTR-Open_Sea-cg04555837 and CENPL-5'UTR-N_Shelf-cg22379576 using MethSurv web tool. Among them, CENPL-3'UTR-Open_Sea-cg04555837 was significantly positively related to shorter survival time, with a HR of 2.431 (Table 7; Figure 7).

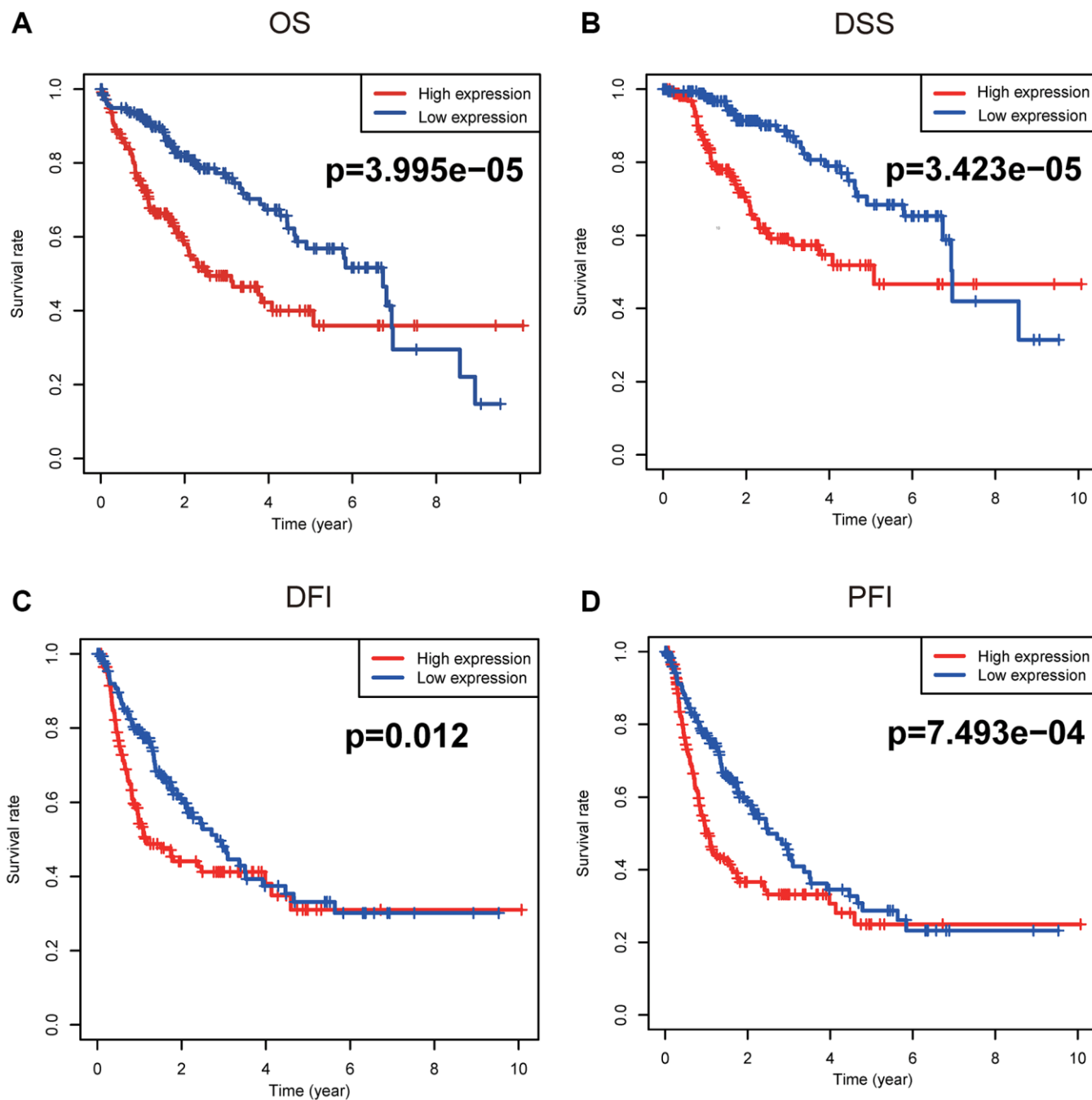


Figure 3. Survival analyses based on Kaplan-Meier method. Kaplan-Meier survival analysis indicated that elevated CENPL was prominently associated with worse (A) OS, (B) DSS, (C) DFI, and (D) PFI. Abbreviations: CENP: centromere protein; OS: overall survival; DSS: disease-specific survival; DFI: disease-free interval; PFI: progression-free interval.

Table 3. Association between clinicopathologic characteristics and OS, DSS, DFI and PFI in patients with HCC through univariate and multivariate analysis with Cox regression survival model.

Characteristics	OS			DSS			DFI			PFI		
	HR	95% CI	P-value	HR	95% CI	P-value	HR	95% CI	P-value	HR	95% CI	P-value
Univariate analysis												
Child-Pugh (B-C vs. A)	1.3	0.4–4.2	0.715	1.8	0.5–6.3	0.363	2.1	0.9–4.9	0.096	1.8	0.8–4.3	0.164
Risk factor (Alcohol consumption and/or viral hepatitis vs. neither)	0.6	0.4–1.0	0.064	0.6	0.3–1.1	0.12	0.9	0.6–1.2	0.341	0.8	0.6–1.1	0.244
AFP (20 ≤ AFP vs. AFP <20)	1.0	0.7–1.6	0.961	1.0	0.6–1.6	0.871	0.9	0.7–1.3	0.632	1.0	0.7–1.3	0.914
New tumor event (yes vs. no)	3.0	1.3–7.0	0.012	1.4	0.9–2.1	0.997	1.0	0.8–1.5	0.996	4.3	1.9–10.1	0.995
Age (>40 vs. ≤40)	4.7	0.6–34.8	0.133	3.8	0.5–28.7	0.197	1.0	0.4–2.3	0.926	0.9	0.4–2.1	0.887
Gender (male vs. female)	0.5	0.2–1.1	0.086	0.6	0.2–1.3	0.195	0.9	0.5–1.6	0.599	0.8	0.5–1.4	0.406
Histologic grade (G3-4 vs. G1-2)	1.6	0.7–3.3	0.247	1.5	0.6–3.6	0.35	1.2	0.7–2.1	0.486	1.1	0.7–1.9	0.645
M (M1 vs. M0)	5.4	0.7–40.7	0.099	8.9	1.1–68.6	0.037	1.1	0.9–1.4	1.000	5.0	0.7–37.4	0.116
N (N1 vs. N0)	4.3	0.6–31.7	0.157	6.1	0.8–46.7	0.081	4.2	0.6–31.4	0.159	3.7	0.5–27.5	0.197
T (T3-4 vs. T1-2)	1.2	0.5–3.0	0.691	1.8	0.7–4.6	0.25	2.0	1.0–3.9	0.043	1.9	1.0–3.6	0.047
Clinical stage (III-IV vs. I-II)	1.4	0.6–3.3	0.454	2.1	0.8–5.2	0.117	2.1	1.1–4.1	0.022	2.0	1.1–3.8	0.026
Postoperative ablation embolization (yes vs. no)	1.1	0.4–3.2	0.87	1.5	0.5–4.6	0.47	2.7	1.4–5.5	0.005	2.7	1.4–5.1	0.003
Radiation therapy (yes vs. no)	0.6	0.3–1.2	0.997	1.3	0.8–2.0	0.997	1.5	0.2–11.1	0.683	1.3	0.2–9.6	0.787
Vascular invasion (yes vs. no)	1.3	0.6–2.9	0.560	1.1	0.4–3.1	0.821	1.0	0.5–2.0	0.910	1.2	0.7–2.1	0.590
Tumor status (with tumor vs. tumor free)	4.0	1.8–9.2	0.001	2.2	1.0–4.9	0.997	35.0	13.4–93.5	0.000	37.0	14.2–97.4	0.000
Family history (yes vs. no)	1.8	0.9–3.7	0.116	2.0	0.8–4.7	0.119	1.1	0.6–2.1	0.639	1.1	0.6–1.9	0.785
Residual tumor (R1-2 vs. R0)	1.4	0.2–10.2	0.761	2.3	0.3–17.4	0.434	1.3	1.1–1.5	0.774	6.4	2.2–18.3	0.001
CENPL (high vs. low)	1.8	1.2–2.6	0.002	2.0	1.3–3.3	0.003	1.9	1.3–2.6	0.000	1.8	1.3–2.5	0.000
Multivariate analysis												
Child-Pugh (B-C vs. A)							3.5	1.3–9.3	0.013			
Risk factor (Alcohol consumption and/or viral hepatitis vs. neither)	0.7	0.4–1.3	0.241									
New tumor event (yes vs. no)	0.4	0.0–3.6	0.412									
Gender (male vs. female)	1.1	0.4–3.1	0.833									
M (M1 vs. M0)	3.2	0.4–26.7	0.285	14.0	1.7–114.3	0.013						
N (N1 vs. N0)				8.9	1.1–70.4	0.039						
T (T3-4 vs. T1-2)							2.4	0.2–22.6	0.458	0.8	0.1–6.3	0.821
Clinical stage (III-IV vs. I-II)							0.7	0.1–6.0	0.706	2.1	0.3–16.5	0.466
Postoperative ablation embolization (yes vs. no)							0.9	0.5–2.0	0.882	1.0	0.5–2.1	0.907
Tumor status (with tumor vs. tumor free)	7.3	0.9–59.6	0.063				39.0	14.3–105.4	0.000	36.0	13.3–96.8	0.000
Residual tumor (R1-2 vs. R0)										2.9	0.9–8.7	0.063
CENPL (high vs. low)	1.7	1.2–2.6	0.008	2.2	1.4–3.6	0.001	1.6	1.1–2.3	0.006	1.7	1.2–2.3	0.002

Abbreviations: OS: overall survival; DSS: disease-specific survival; DFI: disease-free interval; PFI: progression free interval; T: topography distribution; N: lymph node metastasis; M: distant metastasis; HR: hazard ratio; CI: confidence interval; CENPL: centromere protein L.

DISCUSSION

Although breakthroughs have been made in various diagnosis and treatment methods, including biomarkers and immunotherapy, the prognosis of HCC patients is still not optimistic [1]. Therefore, finding markers that can not only predict the occurrence, development and prognosis of HCC, but also predict the effect of

immunotherapy is contributed to correctly diagnosing and intervening HCC in the early stage, improving the prognosis, and reducing unnecessary adverse drug reactions.

Through integrating multiple databases, we found that CNEPL was enhanced in most cancer types, which indicated that CENPL had a wide range of applicability

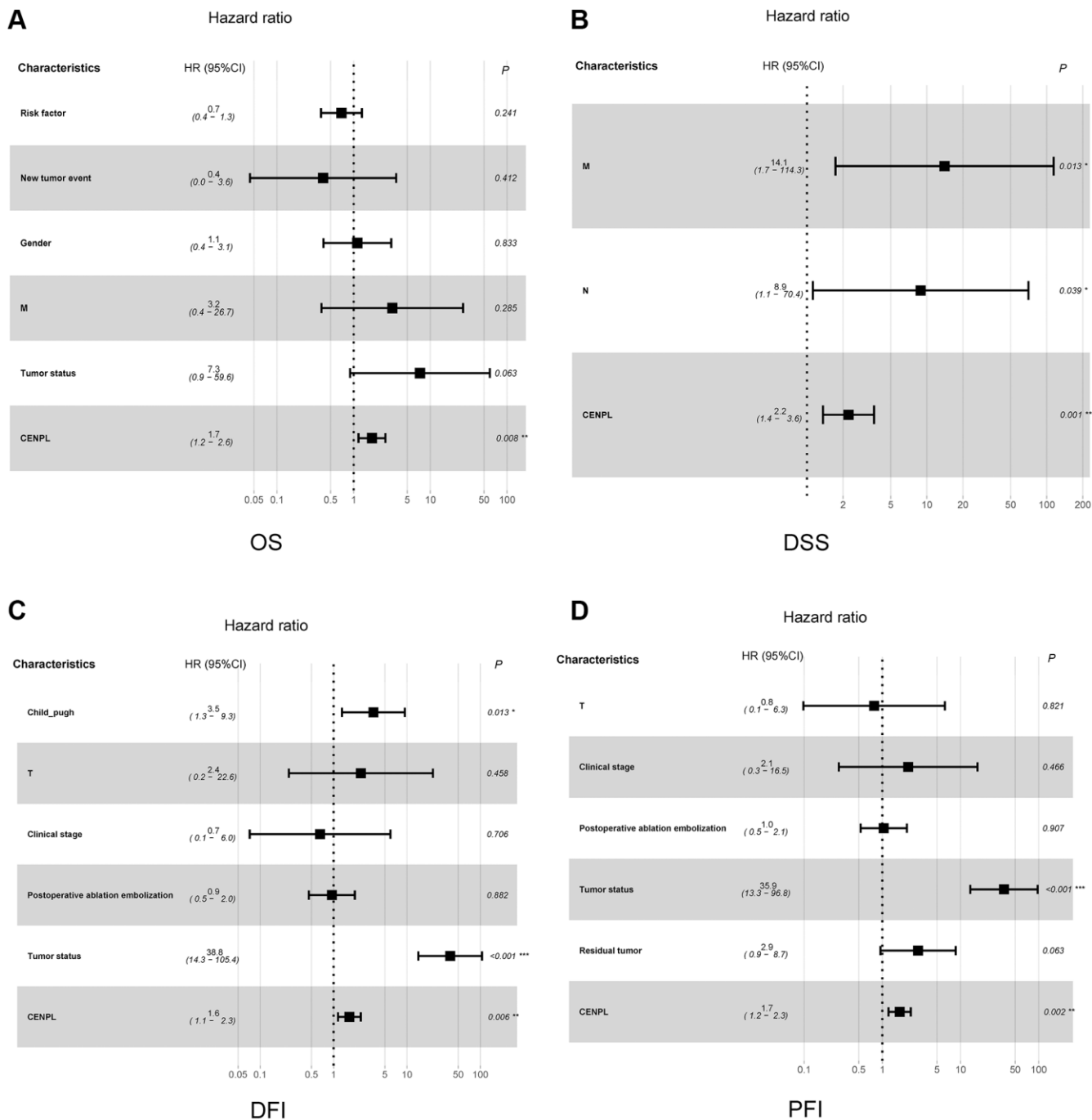


Figure 4. Relationship between clinicopathologic factors and survival outcome of HCC patient through univariate and multivariate Cox regression analysis. (A–D) indicated CENPL may be an independent prognostic factor for poor OS, DSS, DFI and PFI, respectively. Abbreviations: OS: overall survival; DSS: disease-specific survival; DFI: disease-free interval; PFI: progression-free interval.

Table 4. Gene sets enriched in phenotype high.

MSigDB collection	Gene set name	NES	NOM <i>p</i> -value	FDR <i>q</i> -value
c2.cp.kegg.v7.0.symbols.gmt [Curated]	KEGG_UBIQUITIN_MEDIATED_PROTEOLYSIS	2.177	0.000	0.000
	KEGG_CELL_CYCLE	2.132	0.000	0.002
	KEGG_OOCYTE_MEIOSIS	2.128	0.000	0.001
	KEGG_PYRIMIDINE_METABOLISM	2.057	0.000	0.003
	KEGG_NUCLEOTIDE_EXCISION_REPAIR	2.056	0.000	0.002
	KEGG_SPLICEOSOME	1.981	0.002	0.004
	KEGG_DNA_REPLICATION	1.968	0.000	0.004
	KEGG_HOMOLOGOUS_RECOMBINATION	1.940	0.000	0.004
	KEGG_MISMATCH_REPAIR	1.921	0.000	0.004
	KEGG_P53_SIGNALING_PATHWAY	1.884	0.000	0.006

Abbreviations: NES: normalized enrichment score; NOM: nominal; FDR: false discovery rate. Gene sets with NOM *p*-value <0.05 and FDR *q*-value <0.05 are considered as significant.

and functional conservation. However, CENPL was decreased in Acute Myeloid Leukemia (LAML) and testicular germ cell tumor (TGCT) in GEPIA, which indicated that the transcription level of CENPL was still specific in different tumors. We further found that high expression of CENPL was significantly associated with adverse clinicopathological characteristics. Survival analysis revealed that patients with high CENPL expression had worse OS, DSS, DFI, and PFI. Most importantly, univariate and multivariate Cox regression analysis proved that elevated CENPL was an independent risk factor for poor OS, DSS, DFI and PFI in HCC patients. GSEA results indicated that CENPL may be involved in the occurrence and progression of HCC via some pathways. Among these pathways, cell cycle, DNA replication, p53 signaling pathway, and oocyte meiosis play important roles in regulating cell cycle. The hallmarks of cancer comprise sustaining proliferative signaling, evading growth suppressors, resisting cell death, enabling replicative immortality, inducing angiogenesis, and activating invasion and metastasis [25], and these processes all involve cell cycle abnormalities. Nucleotide excision repair [26, 27], mismatch repair [28], homologous recombination [29] are key pathways for repairing DNA damage and preventing tumorigenesis. Disorders in these pathways could lead to genetic mutations, chromosomal aberrations, and subsequent transcriptional and translation errors, which gradually accumulate and result in cancer [26]. Pyrimidine metabolism, a branch of nucleotide metabolism, provides pyrimidine base nucleotides and deoxyribonucleotides to synthesize DNA and RNA, which plays a crucial role in maintaining basic cellular functions [30, 31]. Dysfunctional pyrimidine metabolism can facilitate

cancer proliferation and invasiveness and induce HCC epithelial-mesenchymal transition (EMT), thereby enhancing the stem cell-like characteristics and drug resistance of cancer cells [31–33]. Studies have suggested that pyrimidine metabolism pathways are promising targets for HCC treatment [30, 34]. In addition, spliceosomes exert a key effect in the process of removing introns, connecting exons on both sides and transforming into mature mRNAs after transcription [35–37]. Pre-mRNA splicing is a key step in gene expression [38]. Abnormal RNA spliceosomes and/or splicing processes have been shown to promote tumor genesis and maintenance in a variety of ways [39–41], and some laboratories have begun to develop and design spliceosome inhibitors for anti-tumor effect [42]. Additionally, a meta-analysis reported that the spliceosome pathway was overexpressed in HCC, but its mechanism of action had not been clarified [43]. To our knowledge, this study firstly reported that CENPL may participate in the occurrence and progress of HCC through these pathways.

Co-expressed genes often have similar functions [24, 44], so we explored the co-expression network of CENPL in HCC through the Coexpedia website. We found that Nuf2, also known as Cell Division Cycle Associated 1 (CDCA1), was the strongest positive correlation co-expressed gene. It was reported that Nuf2 was also highly expressed in HCC and played an important role in the arrangement and correct separation of chromosomes during mitosis [45]. Silencing Nuf2 can induce cell cycle arrest, significantly inhibit HCC proliferation and induce cell apoptosis [46]. Then we performed function and pathway enrichment analysis based on the co-expressed genes and found the biological processes (BP) were also mainly related to

cell cycle and metabolism, which were similar to the pathways enriched by CENPL.

With the development of medicine, people gradually realize that traditional TNM staging and pathological

grading can only provide limited prognostic information and cannot predict the response to treatment. More and more researchers begin to pay attention to the host's immune system which plays an important role in controlling tumor occurrence and progression, and

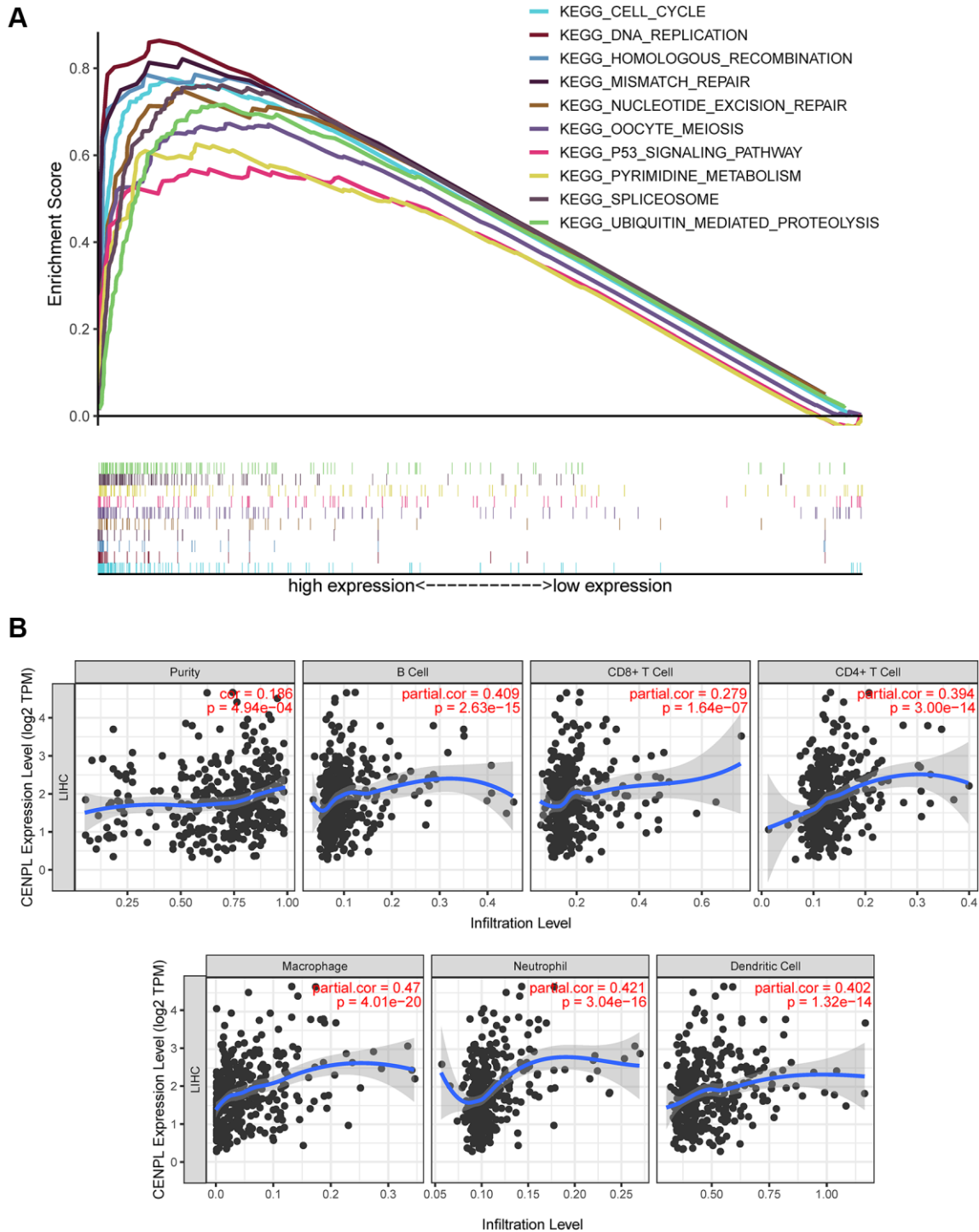


Figure 5. (A) Enrichment plots from GSEA. GSEA results showed that spliceosome, nucleotide excision repair, pyrimidine metabolism, DNA replication, ubiquitin mediated proteolysis, cell cycle, homologous recombination, mismatch repair, p53 signaling pathway and oocyte meiosis were significantly enriched in the high CENPL expression phenotype. GSEA: Gene Set Enrichment Analysis. **(B)** CENPL transcription level had prominent association with level of immune cell infiltration in HCC.

predicting prognosis and therapeutic response [47]. Previous evidences have shown that tumor progression is often closely related to the decrease of CD8+ T cells, NK cells and other potent lymphocytes that play key anti-tumor roles, and the aggregation of regulatory T cells (Tregs) and tumor-associated macrophages [48–51]. Macrophages can stimulate angiogenesis, enhance

tumor cell migration and invasion [52, 53]. This was consistent with the results of our study. Multivariate Cox analysis showed that the infiltration of B cells, CD8+ T cells, and CD4+ T cells were negatively correlated with poor prognosis. That is to say, the decrease of their infiltration indicated an increased risk of poor prognosis. On the contrary,

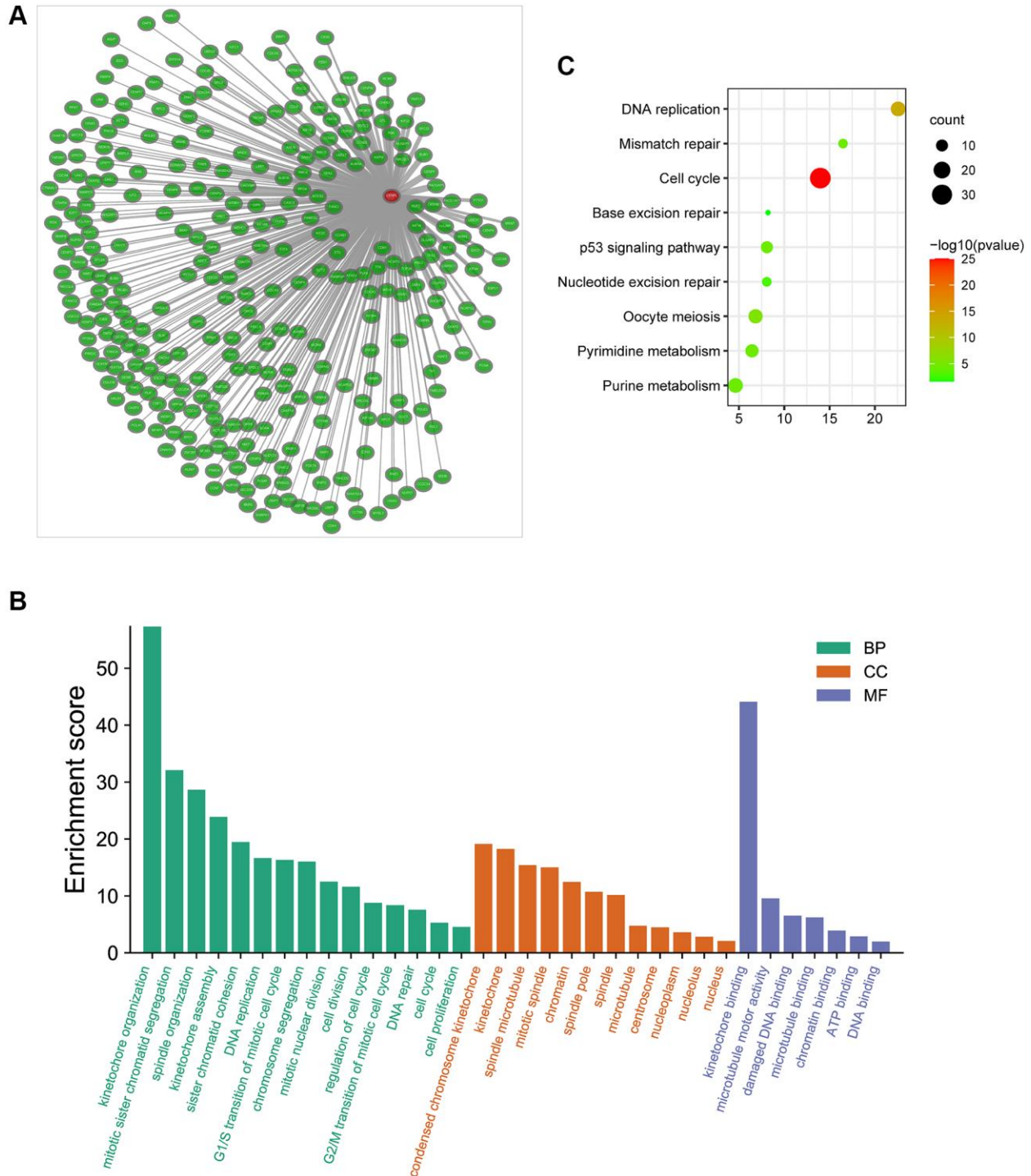


Figure 6. (A) The co-expression network of CENPL; (B–C) GO and KEGG enrichment analysis of co-expression genes of CENPL. Abbreviations: CENPL: centromere protein L; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes.

Table 5. Survival analysis of multivariate COX hazard model based on TIMER online tool.

Clinicopathologic variable	Coef	95% CI	HR	P-value
Age	0.016	1.00–1.03	1.02	0.068
Gender Male	−0.019	0.61–1.59	0.98	0.938
Race Black	1.201	1.20–9.21	3.32	0.021
Race White	−0.002	0.60–1.66	1.00	0.995
Stage II	0.081	0.63–1.87	1.09	0.770
Stage III	0.748	1.30–3.43	2.11	0.003
Stage IV	1.522	1.31–15.99	4.58	0.017
Purity	0.419	0.47–4.95	1.52	0.487
B cell	−7.514	0.00–0.84	0.00	0.045
CD8 T cell	−5.056	0.00–0.92	0.01	0.046
CD4 T cell	−9.241	0.00–0.18	0.00	0.016
Macrophage	8.546	18.24–1452349.07	5146.34	0.003
Neutrophil	−5.479	0.00–478.27	0.00	0.357
Dendritic	4.945	3.02–6526.97	140.42	0.012
CENPL	0.389	1.11–1.96	1.48	0.007

Abbreviations: TIMER: Tumor Immune Estimation Resource; CENPL: centromere protein L; HR: hazard ratio; Coef: coefficient.

Table 6. The association between the expression of CENPL and immune biomarker genes in HCC based on TIMER and GEPIA.

Description	Gene markers	TIMER		GEPIA	
		Cor	P-value	Cor	P-value
CD8+ T cell	CD8A	0.284	8.18E−08	0.160	2.30E−03
	CD8B	0.209	8.84E−05	0.093	7.50E−02
T cell (general)	CD2	0.282	1.03E−07	0.130	1.40E−02
	CD3E	0.282	9.94E−08	0.120	2.30E−02
	CD3D	0.288	4.89E−08	0.120	1.70E−02
	CD6	0.259	1.04E−06	0.130	1.20E−02
	SH2D1A	0.271	3.36E−07	0.140	8.00E−03
	TRAT1	0.263	7.37E−07	0.150	3.40E−03
B cell	CD3G	0.323	7.76E−10	0.250	8.2E−07
	CD19	0.293	2.96E−08	0.230	1.2E−05
Monocyte	CD79A	0.241	5.93E−06	0.120	2.50E−02
	CD86	0.412	1.50E−15	0.270	2.4E−07
TAM	CD115 (CSF1R)	0.295	2.22E−08	0.180	4.90E−04
	CD68	0.282	9.54E−08	0.200	9.9E−05
M1 Macrophage	CCL2	0.216	5.04E−05	0.110	4.30E−02
	IL10	0.284	7.94E−08	0.150	4.60E−03
M2 Macrophage	IRF5	0.478	3.86E−21	0.440	1.1E−18
M2 Macrophage	CD163	0.194	2.91E−04	−0.006	9.10E−01
	VSIG4	0.238	8.12E−06	0.120	2.60E−02

Neutrophil	MS4A4A	0.228	1.99E-05	0.100	5.10E-02
	S100A12	0.000	9.99E-01	-0.023	6.60E-01
	CEACAM3	0.159	3.10E-03	0.120	2.20E-02
	CCR7	0.229	1.69E-05	0.100	5.30E-02
	FPR1	0.322	9.78E-10	0.200	7.4E-05
	SIGLEC5	0.415	7.99E-16	0.240	2.6E-06
	CSF3R	0.375	5.58E-13	0.240	3.9E-06
	FCAR	0.196	2.42E-04	0.094	7.20E-02
	FCGR3B	0.232	1.33E-05	0.320	3.1E-10
NK	KIR2DL1	-0.051	3.41E-01	0.056	2.80E-01
	KIR2DL3	0.253	1.89E-06	0.200	1.50E-04
	KIR2DL4	0.184	5.92E-04	0.200	1E-04
	KIR3DL1	0.039	4.75E-01	-0.048	3.60E-01
	KIR3DL2	0.170	1.54E-03	0.200	7.9E-05
	KIR3DL3	0.186	4.94E-04	0.120	1.90E-02
	XCL1	0.304	8.72E-09	0.240	2.5E-06
	XCL2	0.245	3.98E-06	0.082	1.20E-01
	NCR1	0.119	2.69E-02	0.140	5.90E-03
DC	CD11C (ITGAX)	0.407	3.11E-15	0.300	7.5E-09
	HLA-DPA1	0.280	1.25E-07	0.160	1.70E-03
	HLA-DRA	0.309	4.38E-09	0.190	2.10E-04
	HLA-DQB1	0.204	1.33E-04	0.042	4.20E-01
	HLA-DPB1	0.274	2.23E-07	0.160	1.60E-03
	CCL13	0.162	2.59E-03	0.090	8.30E-02
	HSD11B1	-0.234	1.17E-05	-0.2	9.5E-05
	TBX21 (T-bet)	0.171	1.39E-03	0.072	1.70E-01
Th1	TNF	0.334	1.92E-10	0.230	1.2E-05
	STAT1	0.434	2.76E-17	0.390	7.2E-15
	STAT4	0.296	2.20E-08	0.240	3.1E-06
	IL13	0.107	4.79E-02	0.130	1.10E-02
Th2	GATA3	0.326	5.59E-10	0.210	3.6E-05
	STAT5A	0.386	1.04E-13	0.340	2.1E-11
	STAT6	0.231	1.50E-05	0.280	3.3E-08
	VSIG4	0.241	5.98E-06	0.120	2.60E-02
Th17	STAT3	0.278	1.59E-07	0.270	9.4E-08
Treg	TGFB1	0.366	2.32E-12	0.220	1.6E-05
	STAT5B	0.379	3.29E-13	0.430	2.4E-18
	CCR8	0.504	1.20E-23	0.410	2.7E-16
	FOXP3	0.264	6.30E-07	0.140	5.50E-03
T-cell exhaustion	TIGIT	0.367	1.89E-12	0.230	7.2E-06
	GZMB	0.137	1.10E-02	0.019	7.20E-01
	TOX	0.322	9.49E-10	0.240	5E-06

TIM-3 (HAVCR2)	0.414	9.94E-16	0.250	7.3E-07
LAG3	0.278	1.46E-07	0.140	6.90E-03
CTLA4	0.364	3.02E-12	0.240	2.5E-06
PD-1 (PDCD1)	0.363	3.60E-12	0.240	2.6E-06

Abbreviations: TAM: tumor-associated macrophage; IL: interleukin; Tfh: follicular helper T cell; TNF: tumor necrosis factor; DC: dendritic cell; Treg: regulatory T cell; TIMER: Tumor Immune Estimation Resource; GEPIA: Gene Expression Profiling Interactive Analysis; NK: natural killer cell; Th: T helper cell; HCC: hepatocellular carcinoma; Cor: R value of Spearman's correlation; CENPL: centromere protein L.

Table 7. CpG of CENPL related to the prognosis of HCC.

Gene-CpG	HR	LR test <i>p</i> value
CENPL-3'UTR-Open_Sea-cg04555837	2.431	0.000
CENPL-5'UTR-N_Shelf-cg22379576	1.232	0.232

Abbreviations: CENPL: centromere protein L; HR: hazard ratio; LR: log-rank.

Macrophages were significantly positively associated with poor prognosis, suggesting the high macrophages infiltration predicts adverse prognosis. Elevated CENPL expression could also independently predict worse prognosis under the condition excluding immune cell infiltration. Furthermore, we found that CENPL expression was notably negatively correlated with infiltration of B cells, CD8+ T cells, CD4+ T cells and neutrophil and positively associated with macrophage and dendritic cell, especially had the highest association with macrophages. We speculated that the occurrence and progression of HCC promoted by enhanced CENPL may be partly attributed to augmenting tumorigenic effect of macrophages and attenuating the anti-tumor effect of killer cells such as B cells, CD8+ T cells and CD4+ T cells.

In addition, our results showed that the expression of CENPL was significantly positively correlated with many markers of immune cells, among which IRF5, a marker of M1 macrophages [54], and CCR8 [55] and STAT5B [56], markers of Treg cells, have the highest correlation. Macrophages are often divided into M1 and M2 types according to their functions. Among them, M1 macrophages can secrete inflammatory factors such as interleukin (IL-12), interleukin-6 (IL-6) and tumor necrosis factor alpha (TNF α), promote the production of reactive oxygen species and nitric oxide (NO), and have pro-inflammatory activity [57]. Chronic inflammatory and continuous peroxidation can induce cell cancerization and tumorigenesis. M2 macrophages are involved in accelerating tumor growth, invasion and angiogenesis [52, 53, 57–59]. This study found that the expression of CENPL was positively associated with the markers of M1 and M2 macrophages, but had a higher correlation with M1 macrophages, suggesting that increased CENPL may be more involved in M1

macrophages promoting tumorigenesis. Nevertheless, it may also play a role in the tumor invasion and progression induced by M2 macrophages. Additionally, studies have revealed that Treg cells can inhibit the activity and proliferation of effector CD4+ and CD8+ T cells, and higher Treg cells infiltration means worse prognosis in HCC [49, 50]. This suggests that elevated CENPL may also participate in the process of Treg cells suppressing effector CD4+ and CD8+ T cells. Furthermore, CENPL expression associated prominently and positively with T-cell failure markers such as PD-1, CTLA4, LAG3, TOX, TIGIT, GZMB and Tim-3. Particularly the correlation with Tim-3 was the highest. This explains the result that enhanced CENPL predicts poor prognosis and also provides a basis for finding new immunotherapy methods for HCC patients with poor response to PD-1 antibodies and other immunosuppressive agents.

Tumor immune microenvironment is a complex environment characterized by immunosuppression and immune escape [4, 60]. How to restore the normal anti-tumor immune response to kill tumor cells is a research hotspot in recent years. With the clinical application of immune checkpoint inhibitors (ICIs), new hope has been brought to tumor patients, but only a small number of patients have obtained clinical benefits [4]. Therefore, it is necessary to look for biomarkers to predict the efficacy of ICIs in order to determine the appropriate population for corresponding immunotherapy. Previous studies reported that the expression of PD-L1 and tumor mutational burden (TMB), as biomarkers for the evaluation of ICIs treatment efficacy, could play a better predictive role in some cancers [61–64]. However, other investigators found that PD-L1 and TMB were not significantly associated with ICIs efficacy in most cancer subtypes

[65–67]. Therefore, novel biomarkers which can predict the efficacy of ICIs should be developed in combination with PD-L1 and TMB to correctly assess whether patients can benefit from ICIs treatment [66], thereby improving prognosis and reducing unnecessary drug toxicity in patients who are unlikely to benefit. Previous studies have proved that the levels of major markers of immune cells could represent the abundance of corresponding immune cells and the composition of

immune cell populations in the tumor-immune microenvironment [68], which were helpful to predict the prognosis and select the best immunotherapy scheme [47, 68–71]. Our study revealed that CENPL expression had a prominent correlation with not only the infiltration level of major immune cells, but also major markers of immune cells. This further verifies our conclusion that elevated CENPL can independently predict poor prognosis of HCC, and to some extent

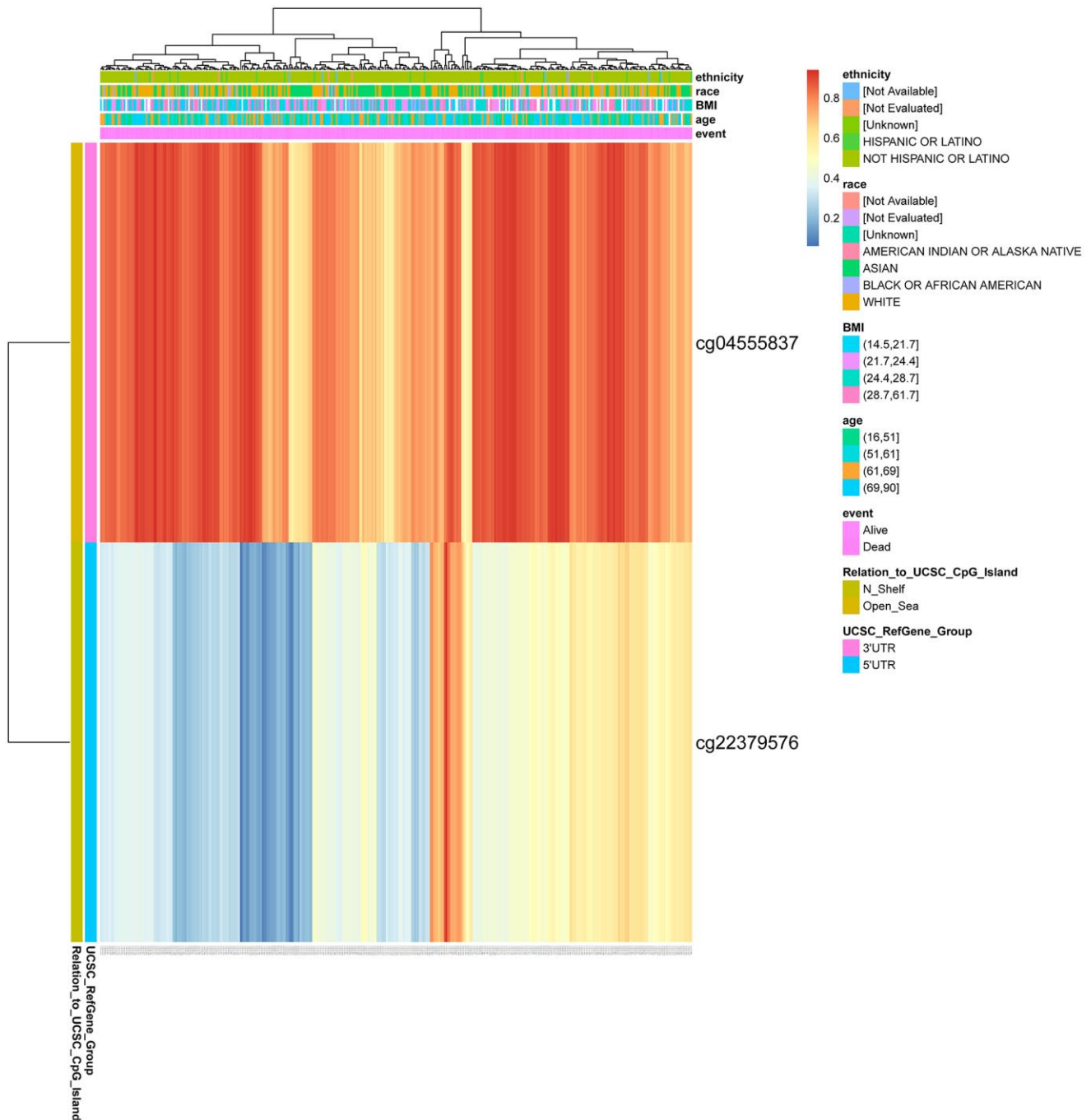


Figure 7. DNA methylation of CENPL. Red to blue means that the expression level goes from high to low. The different colored boxes represent ethnicity, race, BMI, age, event, relation_to_UCSC_CpG_Island, UCSC_RefGene_Group. Abbreviations: CENPL: centromere protein L; BMI: Body Mass Index.

predict the efficacy of some immunotherapy, which provides a direction for new immunotherapy methods.

Many researches provided evidence that almost all tumor types contained abnormal methylation and it could lead to the occurrence of cancer [72, 73]. Studying DNA methylation helps us to understand the mechanism of tumorigenesis and predict the occurrence and progression of cancer [74]. Since methylation is sometimes reversible, it has the potential to become a therapeutic target. Our research unveiled that CENPL had two methylation sites, and CENPL-3'UTR-Open_Sea-cg04555837, one of them, had a significantly positive correlation with shorter survival time.

Although this study is the first to reveal the relationship between CENPL and HCC prognosis and immune infiltration, some false positive rates cannot be ruled out and further experimental and clinical validation is needed.

CONCLUSIONS

Our study uncovered that elevated CENPL in HCC was positively related to adverse clinicopathological factors, occurrence and progression of HCC, and abnormal immunocyte infiltration. It could be an independent predictor for poor prognosis and a promising determinant for immunotherapy.

Abbreviations

CENP: centromere protein; GEO: Gene Expression Omnibus; DEGs: differentially expressed genes; GO: Gene Ontology; KEGG: Kyoto Encyclopedia of Genes and Genomes; HCC: hepatocellular carcinoma; RNA-seq: RNA sequencing; OS: overall survival; DSS: disease-specific survival; DFI: disease-free interval; PFI: progression-free interval.

AUTHOR CONTRIBUTIONS

Z.Z, J.X, Y.T, Z.C and E.Z conceived and designed the study. Z.Z, J.X, Z.P performed the bioinformatics analysis. All authors contributed to data collection and analysis. Z.Z and Z.C prepared figures and tables. Z.Z wrote the manuscript. Z.C, Y.T, Y.C and E.Z reviewed and edited the manuscript. All authors contributed to the manuscript and agreed to be responsible for all aspects of the work.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest related to this study.

FUNDING

This work was supported by Major research projects of Guangzhou University of Chinese Medicine in first-class disciplines (A1-AFD018181A29); Lingnan Traditional Chinese Medicine Tumor Academic School Inheritance Studio Project (No. 2016LP03).

REFERENCES

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin.* 2021; 71:209–49. <https://doi.org/10.3322/caac.21660> PMID:[33538338](https://pubmed.ncbi.nlm.nih.gov/33538338/)
2. Department of Medical Administration, National Health and Health Commission of the People's Republic of China. [Guidelines for diagnosis and treatment of primary liver cancer in China (2019 edition)]. *Zhonghua Gan Zang Bing Za Zhi.* 2020; 28:112–28. <https://doi.org/10.3760/cma.j.issn.1007-3418.2020.02.004> PMID:[32164061](https://pubmed.ncbi.nlm.nih.gov/32164061/)
3. Tiong L, Maddern GJ. Systematic review and meta-analysis of survival and disease recurrence after radiofrequency ablation for hepatocellular carcinoma. *Br J Surg.* 2011; 98:1210–24. <https://doi.org/10.1002/bjs.7669> PMID:[21766289](https://pubmed.ncbi.nlm.nih.gov/21766289/)
4. Binnewies M, Roberts EW, Kersten K, Chan V, Fearon DF, Merad M, Coussens LM, Gabrilovich DI, Ostrand-Rosenberg S, Hedrick CC, Vonderheide RH, Pittet MJ, Jain RK, et al. Understanding the tumor immune microenvironment (TIME) for effective therapy. *Nat Med.* 2018; 24:541–50. <https://doi.org/10.1038/s41591-018-0014-x> PMID:[29686425](https://pubmed.ncbi.nlm.nih.gov/29686425/)
5. Milks KJ, Moree B, Straight AF. Dissection of CENP-C-directed centromere and kinetochore assembly. *Mol Biol Cell.* 2009; 20:4246–55. <https://doi.org/10.1091/mbc.E09-05-0378> PMID:[19641019](https://pubmed.ncbi.nlm.nih.gov/19641019/)
6. Weir JR, Faesen AC, Klare K, Petrovic A, Basilico F, Fischböck J, Pentakota S, Keller J, Pesenti ME, Pan D, Vogt D, Wohlgemuth S, Herzog F, Musacchio A. Insights from biochemical reconstitution into the architecture of human kinetochores. *Nature.* 2016; 537:249–53. <https://doi.org/10.1038/nature19333> PMID:[27580032](https://pubmed.ncbi.nlm.nih.gov/27580032/)

7. Jallepalli PV, Lengauer C. Chromosome segregation and cancer: cutting through the mystery. *Nat Rev Cancer*. 2001; 1:109–17.
<https://doi.org/10.1038/35101065>
PMID:[11905802](https://pubmed.ncbi.nlm.nih.gov/11905802/)
8. Solomon E, Borrow J, Goddard AD. Chromosome aberrations and cancer. *Science*. 1991; 254:1153–60.
<https://doi.org/10.1126/science.1957167>
PMID:[1957167](https://pubmed.ncbi.nlm.nih.gov/1957167/)
9. Guo Q, Tao Y, Liu H, Teng M, Li X. Structural insights into the role of the Chl4-Iml3 complex in kinetochore assembly. *Acta Crystallogr D Biol Crystallogr*. 2013; 69:2412–19.
<https://doi.org/10.1107/S0907444913022397>
PMID:[24311582](https://pubmed.ncbi.nlm.nih.gov/24311582/)
10. McKinley KL, Sekulic N, Guo LY, Tsinman T, Black BE, Cheeseman IM. The CENP-L-N Complex Forms a Critical Node in an Integrated Meshwork of Interactions at the Centromere-Kinetochore Interface. *Mol Cell*. 2015; 60:886–98.
<https://doi.org/10.1016/j.molcel.2015.10.027>
PMID:[26698661](https://pubmed.ncbi.nlm.nih.gov/26698661/)
11. Li Y, Zhu Z, Zhang S, Yu D, Yu H, Liu L, Cao X, Wang L, Gao H, Zhu M. ShRNA-targeted centromere protein A inhibits hepatocellular carcinoma growth. *PLoS One*. 2011; 6:e17794.
<https://doi.org/10.1371/journal.pone.0017794>
PMID:[21423629](https://pubmed.ncbi.nlm.nih.gov/21423629/)
12. Li YM, Liu XH, Cao XZ, Wang L, Zhu MH. [Expression of centromere protein A in hepatocellular carcinoma]. *Zhonghua Bing Li Xue Za Zhi*. 2007; 36:175–78.
PMID:[17535684](https://pubmed.ncbi.nlm.nih.gov/17535684/)
13. Liu Z, Ling K, Wu X, Cao J, Liu B, Li S, Si Q, Cai Y, Yan C, Zhang Y, Weng Y. Reduced expression of cenp-e in human hepatocellular carcinoma. *J Exp Clin Cancer Res*. 2009; 28:156.
<https://doi.org/10.1186/1756-9966-28-156>
PMID:[20021663](https://pubmed.ncbi.nlm.nih.gov/20021663/)
14. Li S, Li X, Xu A, Zhang B, He X, Chen H, Huang J. Screening and clinical evaluation of dominant peptides of centromere protein F antigen for early diagnosis of hepatocellular carcinoma. *Mol Med Rep*. 2018; 17:4720–28.
<https://doi.org/10.3892/mmr.2018.8372>
PMID:[29328419](https://pubmed.ncbi.nlm.nih.gov/29328419/)
15. Uhlén M, Fagerberg L, Hallström BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson Å, Kampf C, Sjöstedt E, Asplund A, Olsson I, Edlund K, Lundberg E, et al. Proteomics. Tissue-based map of the human proteome. *Science*. 2015; 347:1260419.
<https://doi.org/10.1126/science.1260419>
PMID:[25613900](https://pubmed.ncbi.nlm.nih.gov/25613900/)
16. Thul PJ, Åkesson L, Wiking M, Mahdessian D, Geladaki A, Ait Blal H, Alm T, Asplund A, Björk L, Breckels LM, Bäckström A, Danielsson F, Fagerberg L, et al. A subcellular map of the human proteome. *Science*. 2017; 356:eaal3321.
<https://doi.org/10.1126/science.aal3321>
PMID:[28495876](https://pubmed.ncbi.nlm.nih.gov/28495876/)
17. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005; 102:15545–50.
<https://doi.org/10.1073/pnas.0506580102>
PMID:[16199517](https://pubmed.ncbi.nlm.nih.gov/16199517/)
18. Nagy Á, Munkácsy G, Györfly B. Pancancer survival analysis of cancer hallmark genes. *Sci Rep*. 2021; 11:6047.
<https://doi.org/10.1038/s41598-021-84787-5>
PMID:[33723286](https://pubmed.ncbi.nlm.nih.gov/33723286/)
19. Yang S, Kim CY, Hwang S, Kim E, Kim H, Shim H, Lee I. COEXPEDIA: exploring biomedical hypotheses via co-expressions associated with medical subject headings (MeSH). *Nucleic Acids Res*. 2017; 45:D389–96.
<https://doi.org/10.1093/nar/gkw868>
PMID:[27679477](https://pubmed.ncbi.nlm.nih.gov/27679477/)
20. Huang DW, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc*. 2009; 4:44–57.
<https://doi.org/10.1038/nprot.2008.211>
PMID:[19131956](https://pubmed.ncbi.nlm.nih.gov/19131956/)
21. Li T, Fan J, Wang B, Traugh N, Chen Q, Liu JS, Li B, Liu XS. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. *Cancer Res*. 2017; 77:e108–10.
<https://doi.org/10.1158/0008-5472.CAN-17-0307>
PMID:[29092952](https://pubmed.ncbi.nlm.nih.gov/29092952/)
22. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. 2017; 45:W98–102.
<https://doi.org/10.1093/nar/gkx247>
PMID:[28407145](https://pubmed.ncbi.nlm.nih.gov/28407145/)
23. Modhukur V, Ilijasenکو T, Metsalu T, Lökk K, Laisk-Podar T, Vilo J. MethSurv: a web tool to perform multivariable survival analysis using DNA methylation data. *Epigenomics*. 2018; 10:277–88.
<https://doi.org/10.2217/epi-2017-0118>
PMID:[29264942](https://pubmed.ncbi.nlm.nih.gov/29264942/)

24. Kakati T, Bhattacharyya DK, Barah P, Kalita JK. Comparison of Methods for Differential Co-expression Analysis for Disease Biomarker Prediction. *Comput Biol Med.* 2019; 113:103380. <https://doi.org/10.1016/j.combiomed.2019.103380> PMID:31415946
25. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. *Cell.* 2011; 144:646–74. <https://doi.org/10.1016/j.cell.2011.02.013> PMID:21376230
26. Marteiijn JA, Lans H, Vermeulen W, Hoeijmakers JH. Understanding nucleotide excision repair and its roles in cancer and ageing. *Nat Rev Mol Cell Biol.* 2014; 15:465–81. <https://doi.org/10.1038/nrm3822> PMID:24954209
27. Friedberg EC. How nucleotide excision repair protects against cancer. *Nat Rev Cancer.* 2001; 1:22–33. <https://doi.org/10.1038/35094000> PMID:11900249
28. Peltomäki P. Role of DNA mismatch repair defects in the pathogenesis of human cancer. *J Clin Oncol.* 2003; 21:1174–79. <https://doi.org/10.1200/JCO.2003.04.060> PMID:12637487
29. Evers B, Helleday T, Jonkers J. Targeting homologous recombination repair defects in cancer. *Trends Pharmacol Sci.* 2010; 31:372–80. <https://doi.org/10.1016/j.tips.2010.06.001> PMID:20598756
30. Wang W, Cui J, Ma H, Lu W, Huang J. Targeting Pyrimidine Metabolism in the Era of Precision Cancer Medicine. *Front Oncol.* 2021; 11:684961. <https://doi.org/10.3389/fonc.2021.684961> PMID:34123854
31. Siddiqui A, Ceppi P. A non-proliferative role of pyrimidine metabolism in cancer. *Mol Metab.* 2020; 35:100962. <https://doi.org/10.1016/j.molmet.2020.02.005> PMID:32244187
32. Bhowmik SK, Ramirez-Peña E, Arnold JM, Putluri V, Sphyris N, Michailidis G, Putluri N, Ambs S, Sreekumar A, Mani SA. EMT-induced metabolite signature identifies poor clinical outcome. *Oncotarget.* 2015; 6:42651–60. <https://doi.org/10.18632/oncotarget.4765> PMID:26315396
33. Garavito MF, Narváez-Ortiz HY, Zimmermann BH. Pyrimidine Metabolism: Dynamic and Versatile Pathways in Pathogens and Cellular Development. *J Genet Genomics.* 2015; 42:195–205. <https://doi.org/10.1016/j.jgg.2015.04.004> PMID:26059768
34. Yeh HW, Lee SS, Chang CY, Hu CM, Jou YS. Pyrimidine metabolic rate limiting enzymes in poorly-differentiated hepatocellular carcinoma are signature genes of cancer stemness and associated with poor prognosis. *Oncotarget.* 2017; 8:77734–51. <https://doi.org/10.18632/oncotarget.20774> PMID:29100421
35. Bai R, Yan C, Wan R, Lei J, Shi Y. Structure of the Post-catalytic Spliceosome from *Saccharomyces cerevisiae*. *Cell.* 2017; 171:1589–98.e8. <https://doi.org/10.1016/j.cell.2017.10.038> PMID:29153833
36. Yan C, Hang J, Wan R, Huang M, Wong CC, Shi Y. Structure of a yeast spliceosome at 3.6-angstrom resolution. *Science.* 2015; 349:1182–91. <https://doi.org/10.1126/science.aac7629> PMID:26292707
37. Wan R, Yan C, Bai R, Huang G, Shi Y. Structure of a yeast catalytic step I spliceosome at 3.4 Å resolution. *Science.* 2016; 353:895–904. <https://doi.org/10.1126/science.aag2235> PMID:27445308
38. Maniatis T, Tasic B. Alternative pre-mRNA splicing and proteome expansion in metazoans. *Nature.* 2002; 418:236–43. <https://doi.org/10.1038/418236a> PMID:12110900
39. Chabot B, Shkreta L. Defective control of pre-messenger RNA splicing in human disease. *J Cell Biol.* 2016; 212:13–27. <https://doi.org/10.1083/jcb.201510032> PMID:26728853
40. Dvinge H, Kim E, Abdel-Wahab O, Bradley RK. RNA splicing factors as oncoproteins and tumour suppressors. *Nat Rev Cancer.* 2016; 16:413–30. <https://doi.org/10.1038/nrc.2016.51> PMID:27282250
41. Blijlevens M, Li J, van Beusechem VW. Biology of the mRNA Splicing Machinery and Its Dysregulation in Cancer Providing Therapeutic Opportunities. *Int J Mol Sci.* 2021; 22:5110. <https://doi.org/10.3390/ijms22105110> PMID:34065983
42. Eymin B. Targeting the spliceosome machinery: A new therapeutic axis in cancer? *Biochem Pharmacol.* 2021; 189:114039. <https://doi.org/10.1016/j.bcp.2020.114039> PMID:32417188
43. Xu W, Huang H, Yu L, Cao L. Meta-analysis of gene expression profiles indicates genes in spliceosome

- pathway are up-regulated in hepatocellular carcinoma (HCC). *Med Oncol.* 2015; 32:96.
<https://doi.org/10.1007/s12032-014-0425-6>
PMID:[25731616](https://pubmed.ncbi.nlm.nih.gov/25731616/)
44. Xie J, Ma A, Fennell A, Ma Q, Zhao J. It is time to apply biclustering: a comprehensive review of biclustering applications in biological and biomedical data. *Brief Bioinform.* 2019; 20:1449–64.
<https://doi.org/10.1093/bib/bby014>
PMID:[29490019](https://pubmed.ncbi.nlm.nih.gov/29490019/)
45. Hu P, Chen X, Sun J, Bie P, Zhang LD. siRNA-mediated knockdown against NUF2 suppresses pancreatic cancer proliferation in vitro and in vivo. *Biosci Rep.* 2015; 35:e00170.
<https://doi.org/10.1042/BSR20140124>
PMID:[25370920](https://pubmed.ncbi.nlm.nih.gov/25370920/)
46. Liu Q, Dai SJ, Li H, Dong L, Peng YP. Silencing of NUF2 inhibits tumor growth and induces apoptosis in human hepatocellular carcinomas. *Asian Pac J Cancer Prev.* 2014; 15:8623–29.
<https://doi.org/10.7314/apjcp.2014.15.20.8623>
PMID:[25374179](https://pubmed.ncbi.nlm.nih.gov/25374179/)
47. Galon J, Pagès F, Marincola FM, Angell HK, Thurin M, Lugli A, Zlobec I, Berger A, Bifulco C, Botti G, Tatangelo F, Britten CM, Kreiter S, et al. Cancer classification using the Immunoscore: a worldwide task force. *J Transl Med.* 2012; 10:205.
<https://doi.org/10.1186/1479-5876-10-205>
PMID:[23034130](https://pubmed.ncbi.nlm.nih.gov/23034130/)
48. Palucka AK, Coussens LM. The Basis of Oncoimmunology. *Cell.* 2016; 164:1233–47.
<https://doi.org/10.1016/j.cell.2016.01.049>
PMID:[26967289](https://pubmed.ncbi.nlm.nih.gov/26967289/)
49. Unitt E, Rushbrook SM, Marshall A, Davies S, Gibbs P, Morris LS, Coleman N, Alexander GJ. Compromised lymphocytes infiltrate hepatocellular carcinoma: the role of T-regulatory cells. *Hepatology.* 2005; 41:722–30.
<https://doi.org/10.1002/hep.20644>
PMID:[15791620](https://pubmed.ncbi.nlm.nih.gov/15791620/)
50. Mathai AM, Kapadia MJ, Alexander J, Kernochan LE, Swanson PE, Yeh MM. Role of Foxp3-positive tumor-infiltrating lymphocytes in the histologic features and clinical outcomes of hepatocellular carcinoma. *Am J Surg Pathol.* 2012; 36:980–86.
<https://doi.org/10.1097/PAS.0b013e31824e9b7c>
PMID:[22446942](https://pubmed.ncbi.nlm.nih.gov/22446942/)
51. Kurebayashi Y, Ojima H, Tsujikawa H, Kubota N, Maehara J, Abe Y, Kitago M, Shinoda M, Kitagawa Y, Sakamoto M. Landscape of immune microenvironment in hepatocellular carcinoma and its additional impact on histological and molecular classification. *Hepatology.* 2018; 68:1025–41.
<https://doi.org/10.1002/hep.29904>
PMID:[29603348](https://pubmed.ncbi.nlm.nih.gov/29603348/)
52. Wynn TA, Chawla A, Pollard JW. Macrophage biology in development, homeostasis and disease. *Nature.* 2013; 496:445–55.
<https://doi.org/10.1038/nature12034>
PMID:[23619691](https://pubmed.ncbi.nlm.nih.gov/23619691/)
53. Qian BZ, Pollard JW. Macrophage diversity enhances tumor progression and metastasis. *Cell.* 2010; 141:39–51.
<https://doi.org/10.1016/j.cell.2010.03.014>
PMID:[20371344](https://pubmed.ncbi.nlm.nih.gov/20371344/)
54. Moradi-Chaleshtori M, Bandehpour M, Heidari N, Mohammadi-Yeganeh S, Mahmoud Hashemi S. Exosome-mediated miR-33 transfer induces M1 polarization in mouse macrophages and exerts antitumor effect in 4T1 breast cancer cell line. *Int Immunopharmacol.* 2021; 90:107198.
<https://doi.org/10.1016/j.intimp.2020.107198>
PMID:[33249048](https://pubmed.ncbi.nlm.nih.gov/33249048/)
55. Chakraborty S, Zappasodi R. To Go or Not to Go?- Targeting Tregs Traveling in Tumors. *Cancer Res.* 2021; 81:2817–19.
<https://doi.org/10.1158/0008-5472.CAN-21-1203>
PMID:[34087782](https://pubmed.ncbi.nlm.nih.gov/34087782/)
56. Wongkajornsilp A, Numchaisermasuk N, Sa-Ngiamsunton K, Akarasereenont P, Wamanuttajinda V, Kasetsinsombat K, Duangsa-Ard S, Laohapan T, Maneechotesuwan K. Effects of the Ayurved Siriraj Wattana recipe on functional and phenotypic characterization of cytokine-induced killer cells and dendritic cells in vitro. *BMC Complement Altern Med.* 2016; 16:489.
<https://doi.org/10.1186/s12906-016-1480-7>
PMID:[27899095](https://pubmed.ncbi.nlm.nih.gov/27899095/)
57. Wu T, Dai Y. Tumor microenvironment and therapeutic response. *Cancer Lett.* 2017; 387:61–68.
<https://doi.org/10.1016/j.canlet.2016.01.043>
PMID:[26845449](https://pubmed.ncbi.nlm.nih.gov/26845449/)
58. Jayasingam SD, Citartan M, Thang TH, Mat Zin AA, Ang KC, Ch'ng ES. Evaluating the Polarization of Tumor-Associated Macrophages Into M1 and M2 Phenotypes in Human Cancer Tissue: Technicalities and Challenges in Routine Clinical Practice. *Front Oncol.* 2020; 9:1512.
<https://doi.org/10.3389/fonc.2019.01512>
PMID:[32039007](https://pubmed.ncbi.nlm.nih.gov/32039007/)
59. Aras S, Zaidi MR. TAMEless traitors: macrophages in cancer progression and metastasis. *Br J Cancer.* 2017; 117:1583–91.
<https://doi.org/10.1038/bjc.2017.356>
PMID:[29065107](https://pubmed.ncbi.nlm.nih.gov/29065107/)

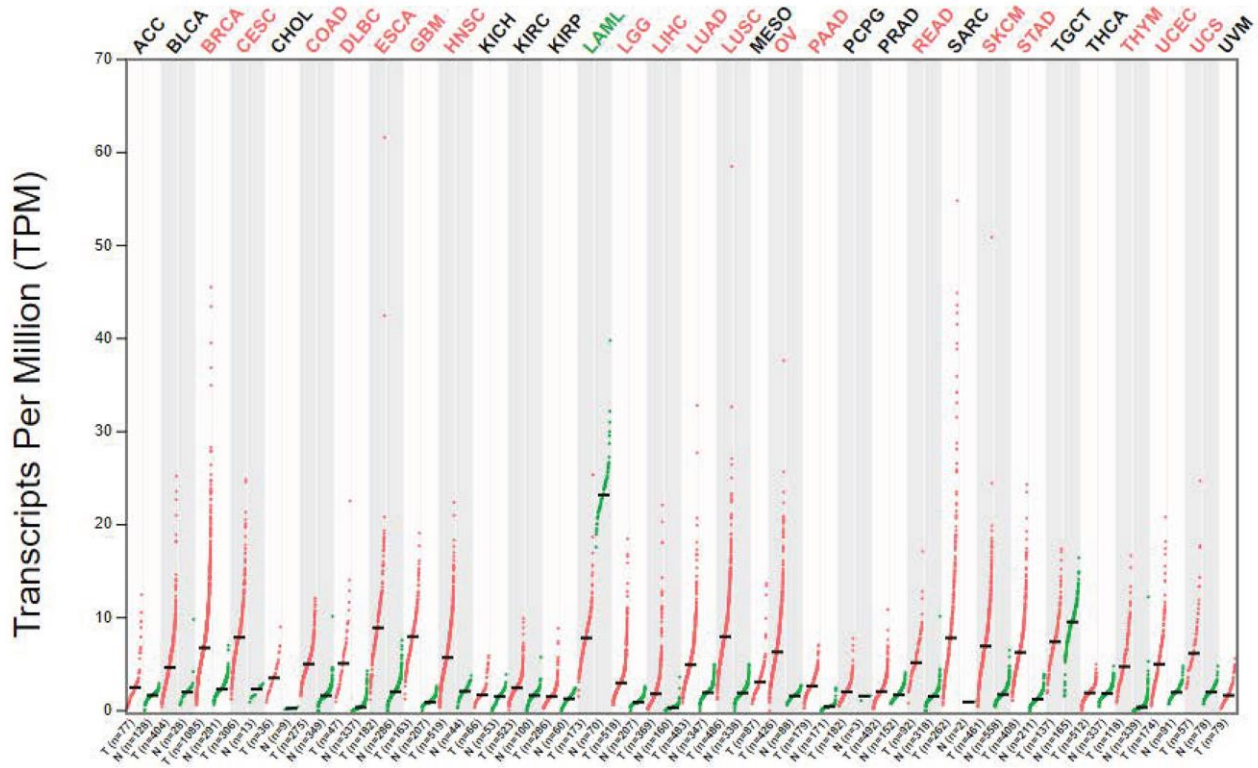
60. Saeed M, Chen F, Ye J, Shi Y, Lammers T, De Geest BG, Xu ZP, Yu H. From Design to Clinic: Engineered Nanobiomaterials for Immune Normalization Therapy of Cancer. *Adv Mater.* 2021; 33:e2008094. <https://doi.org/10.1002/adma.202008094> PMID:34048101
61. Patel SP, Kurzrock R. PD-L1 Expression as a Predictive Biomarker in Cancer Immunotherapy. *Mol Cancer Ther.* 2015; 14:847–56. <https://doi.org/10.1158/1535-7163.MCT-14-0983> PMID:25695955
62. Le DT, Durham JN, Smith KN, Wang H, Bartlett BR, Aulakh LK, Lu S, Kemberling H, Wilt C, Luber BS, Wong F, Azad NS, Rucki AA, et al. Mismatch repair deficiency predicts response of solid tumors to PD-1 blockade. *Science.* 2017; 357:409–13. <https://doi.org/10.1126/science.aan6733> PMID:28596308
63. Le DT, Uram JN, Wang H, Bartlett BR, Kemberling H, Eyring AD, Skora AD, Luber BS, Azad NS, Laheru D, Biedrzycki B, Donehower RC, Zaheer A, et al. PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *N Engl J Med.* 2015; 372:2509–20. <https://doi.org/10.1056/NEJMoa1500596> PMID:26028255
64. Yarchoan M, Hopkins A, Jaffee EM. Tumor Mutational Burden and Response Rate to PD-1 Inhibition. *N Engl J Med.* 2017; 377:2500–01. <https://doi.org/10.1056/NEJMc1713444> PMID:29262275
65. Cristescu R, Mogg R, Ayers M, Albright A, Murphy E, Yearley J, Sher X, Liu XQ, Lu H, Nebozhyn M, Zhang C, Lunceford JK, Joe A, et al. Pan-tumor genomic biomarkers for PD-1 checkpoint blockade-based immunotherapy. *Science.* 2018; 362:eaar3593. <https://doi.org/10.1126/science.aar3593> PMID:30309915
66. Jardim DL, Goodman A, de Melo Gagliato D, Kurzrock R. The Challenges of Tumor Mutational Burden as an Immunotherapy Biomarker. *Cancer Cell.* 2021; 39:154–73. <https://doi.org/10.1016/j.ccell.2020.10.001> PMID:33125859
67. Yarchoan M, Albacker LA, Hopkins AC, Montesion M, Murugesan K, Vithayathil TT, Zaidi N, Azad NS, Laheru DA, Frampton GM, Jaffee EM. PD-L1 expression and tumor mutational burden are independent biomarkers in most cancers. *JCI Insight.* 2019; 4:e126908. <https://doi.org/10.1172/jci.insight.126908> PMID:30895946
68. Danaher P, Warren S, Dennis L, D'Amico L, White A, Disis ML, Geller MA, Odunsi K, Beechem J, Fling SP. Gene expression markers of Tumor Infiltrating Leukocytes. *J Immunother Cancer.* 2017; 5:18. <https://doi.org/10.1186/s40425-017-0215-8> PMID:28239471
69. Jochems C, Schlom J. Tumor-infiltrating immune cells and prognosis: the potential link between conventional cancer therapy and immunity. *Exp Biol Med (Maywood).* 2011; 236:567–79. <https://doi.org/10.1258/ebm.2011.011007> PMID:21486861
70. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, Nair VS, Xu Y, Khuong A, Hoang CD, Diehn M, West RB, Plevritis SK, Alizadeh AA. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med.* 2015; 21:938–45. <https://doi.org/10.1038/nm.3909> PMID:26193342
71. Fridman WH, Pagès F, Sautès-Fridman C, Galon J. The immune contexture in human tumours: impact on clinical outcome. *Nat Rev Cancer.* 2012; 12:298–306. <https://doi.org/10.1038/nrc3245> PMID:22419253
72. Yu DH, Waterland RA, Zhang P, Schady D, Chen MH, Guan Y, Gadkari M, Shen L. Targeted p16(Ink4a) epimutation causes tumorigenesis and reduces survival in mice. *J Clin Invest.* 2014; 124:3708–12. <https://doi.org/10.1172/JCI76507> PMID:25061879
73. Chen SS, Raval A, Johnson AJ, Hertlein E, Liu TH, Jin VX, Sherman MH, Liu SJ, Dawson DW, Williams KE, Lanasa M, Liyanarachchi S, Lin TS, et al. Epigenetic changes during disease progression in a murine model of human chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A.* 2009; 106:13433–38. <https://doi.org/10.1073/pnas.0906455106> PMID:19666576
74. Legendre C, Gooden GC, Johnson K, Martinez RA, Liang WS, Salhia B. Whole-genome bisulfite sequencing of cell-free DNA identifies signature associated with metastatic breast cancer. *Clin Epigenetics.* 2015; 7:100. <https://doi.org/10.1186/s13148-015-0135-8> PMID:26380585

SUPPLEMENTARY MATERIALS

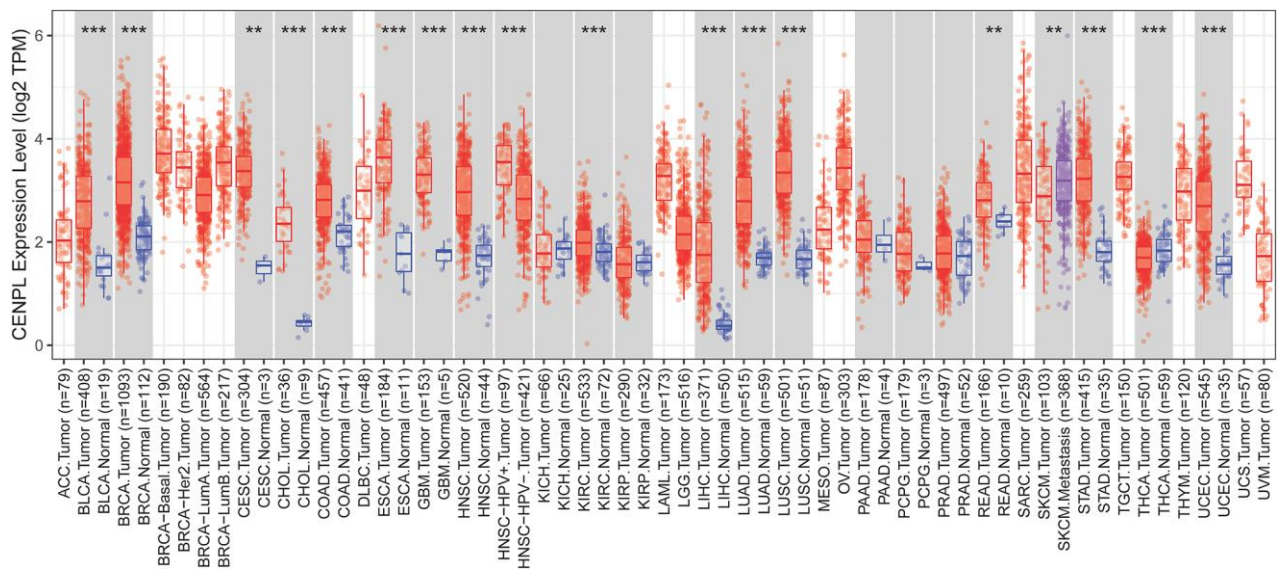
Supplementary Figures

The gene expression profile across all tumor samples and paired normal tissues.(Dot plot)

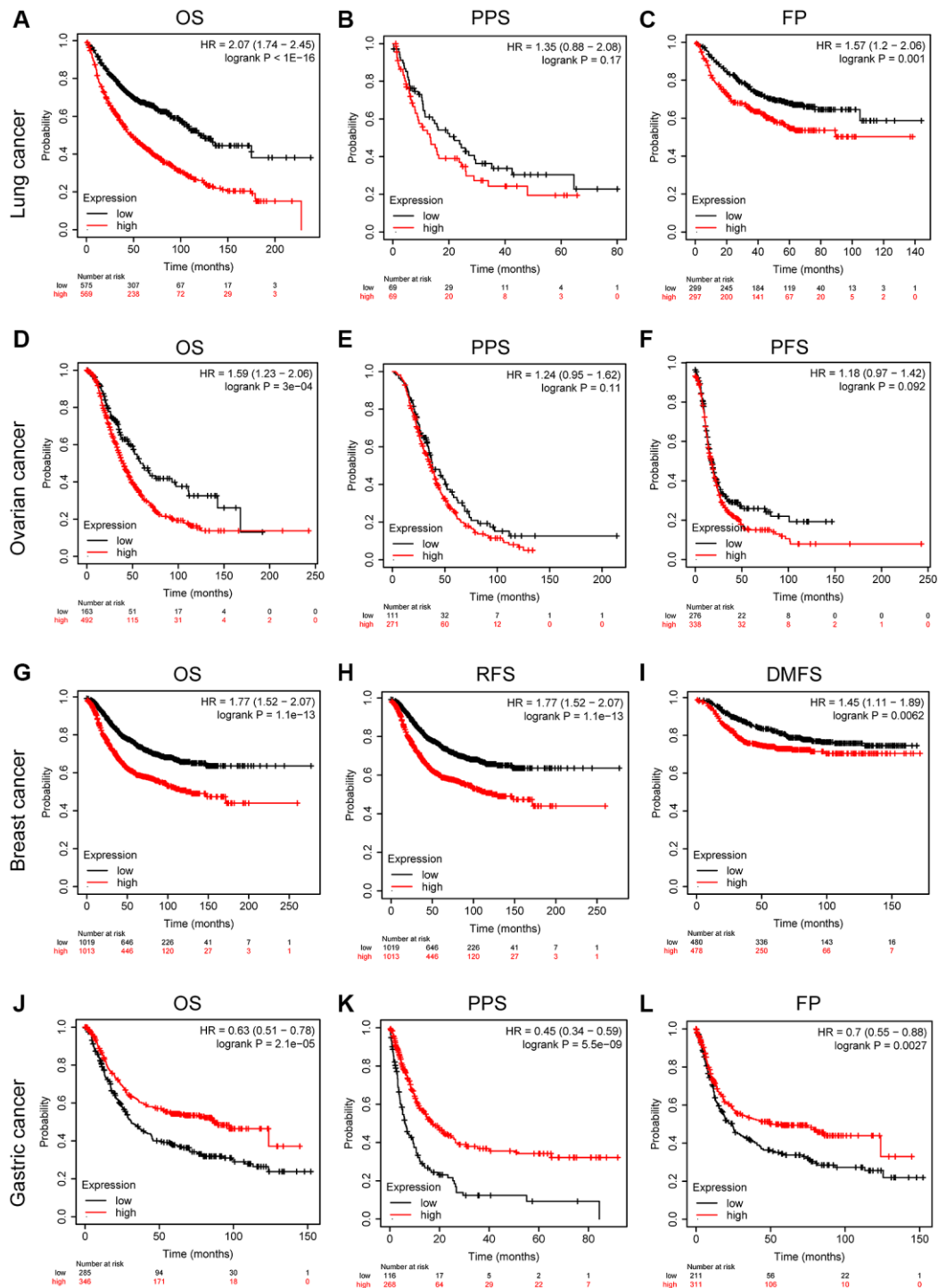
Each dots represent expression of samples.



Supplementary Figure 1. The transcription level of CENPL in different tumor types by GEPIA.



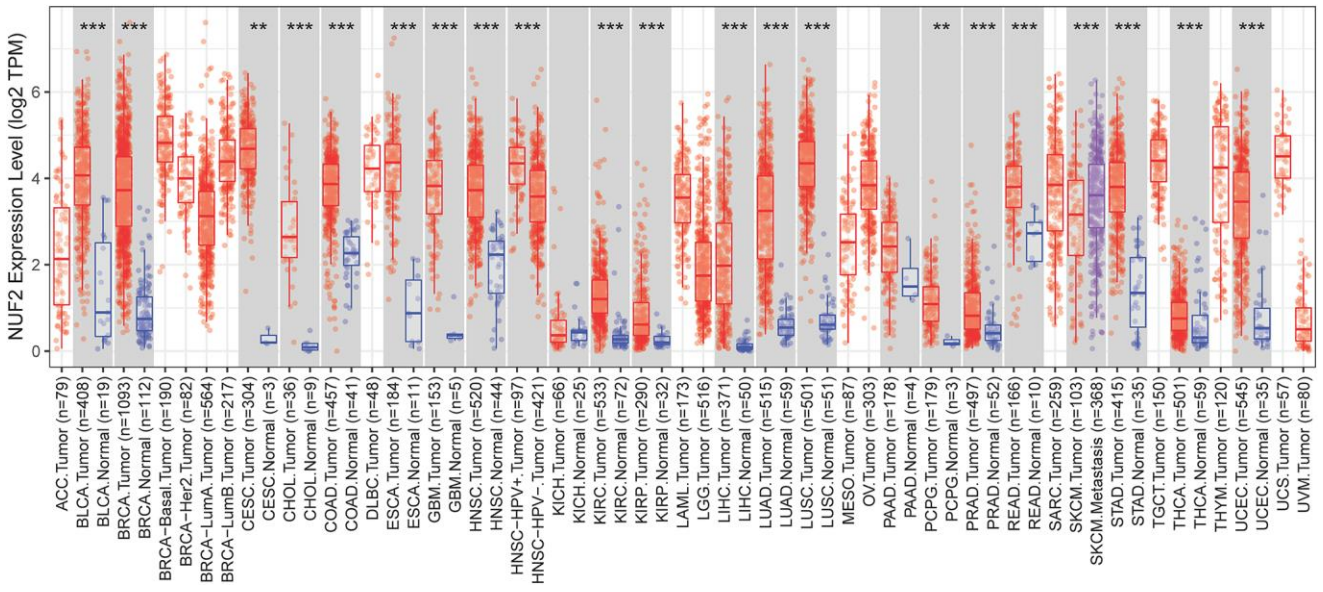
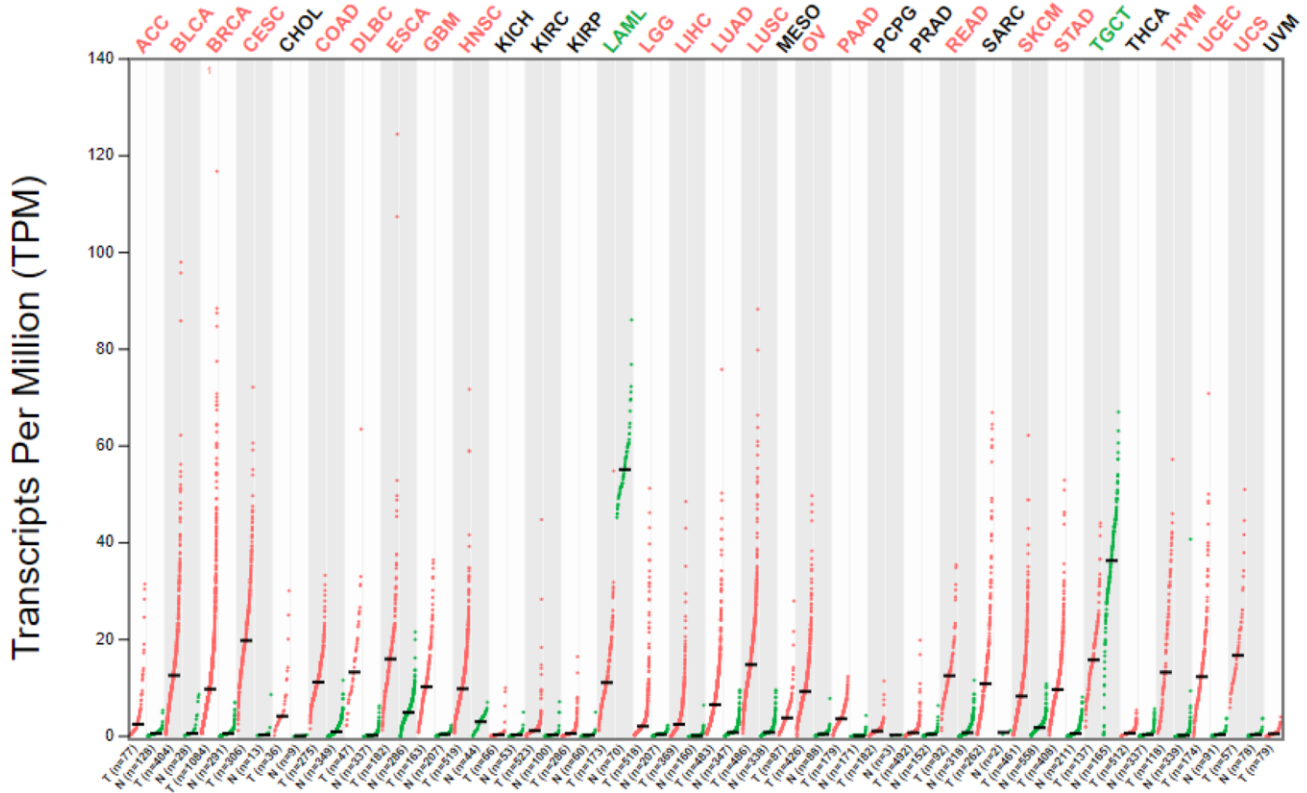
Supplementary Figure 2. The transcription level of CENPL in different tumor types through TIMER.



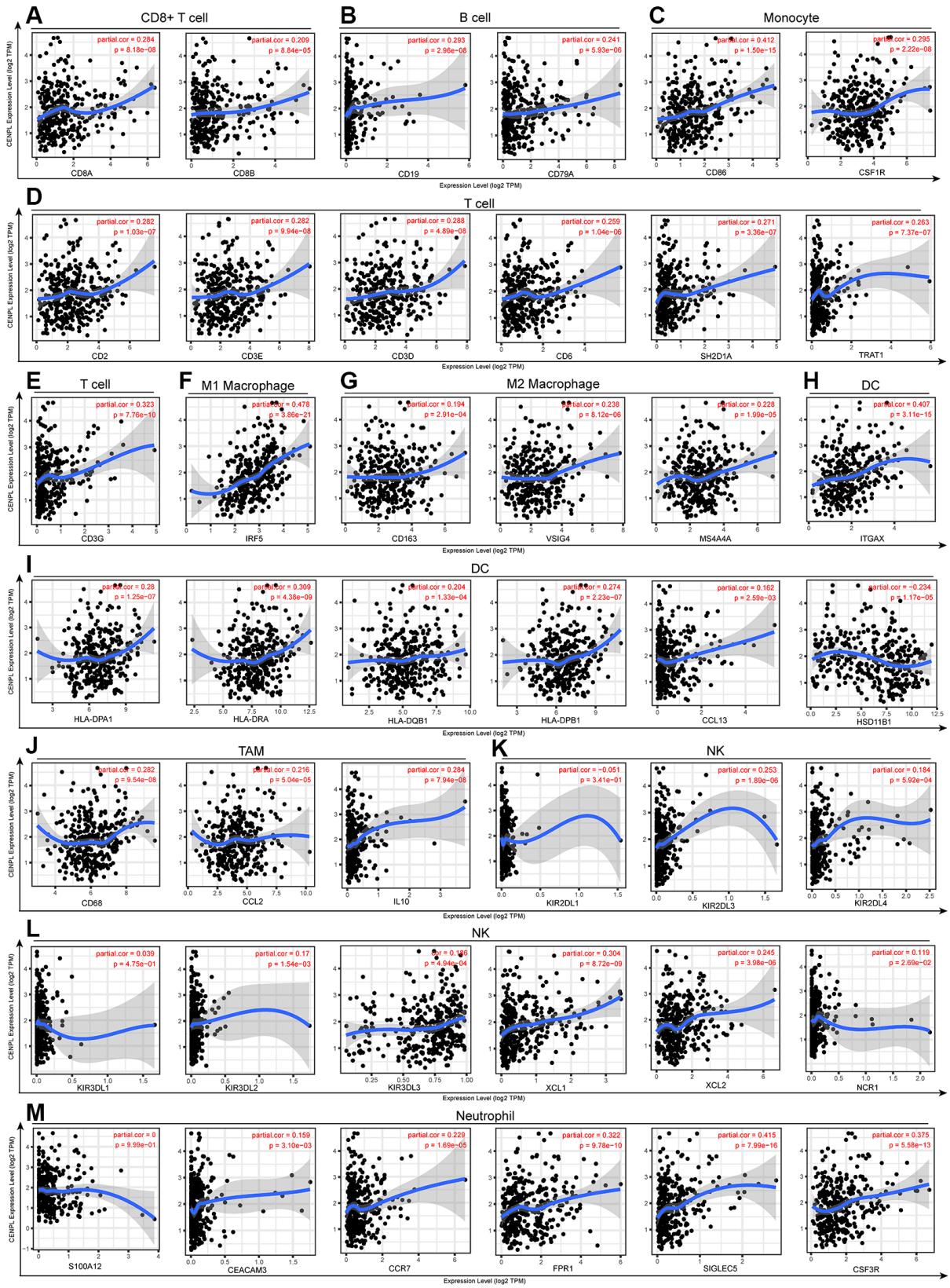
Supplementary Figure 3. The effects of CENPL expression on the prognostic outcome of other tumor types. (A–C) Lung cancer; (D–F) Ovarian cancer; (G–I) Breast cancer, (J–L) Gastric cancer. Abbreviations: OS: overall survival; PPS: post progression survival; FP: first progression; PFS: progression-free survival; RFS: relapse-free survival; DMFS: distance metastasis free survival.

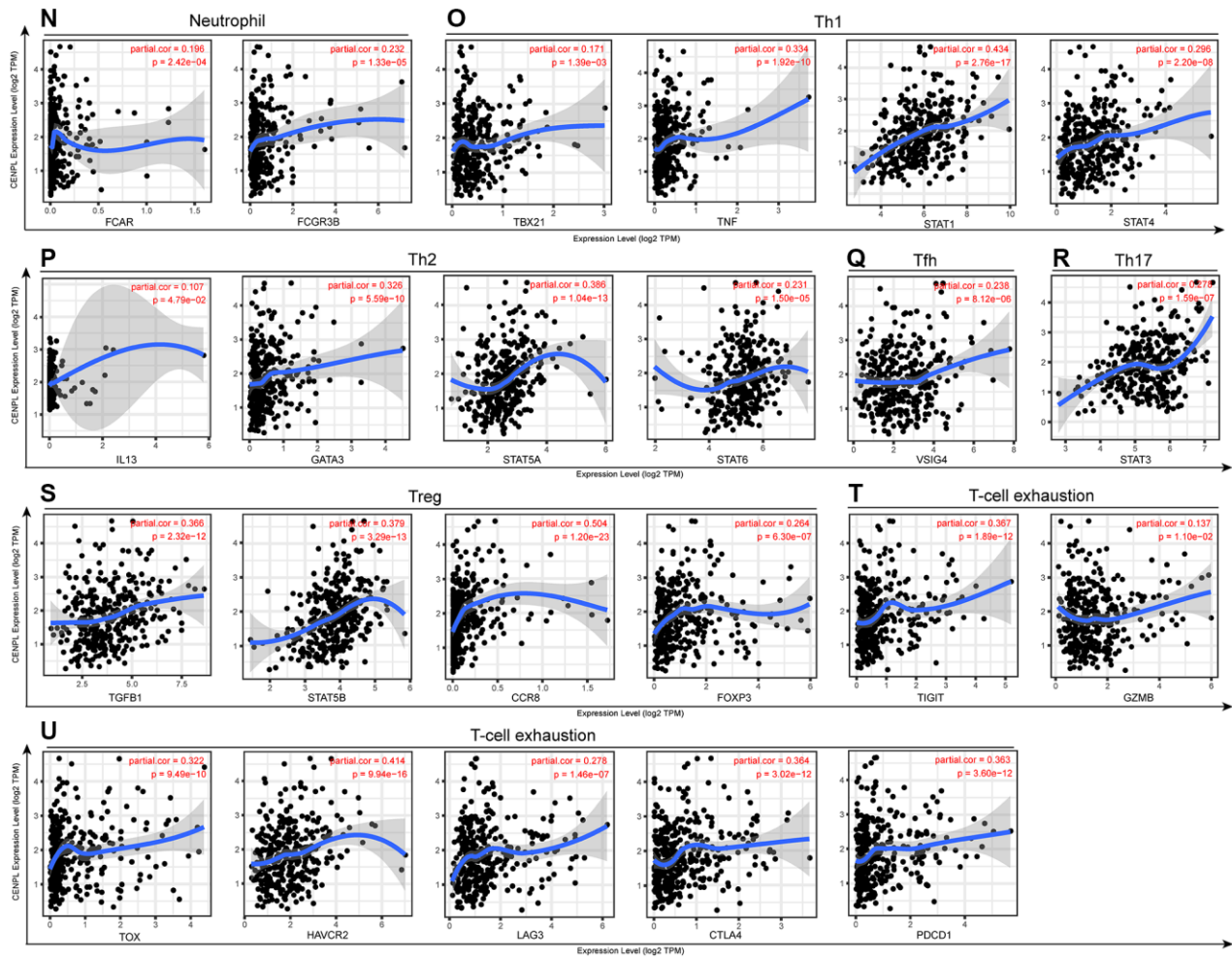
The gene expression profile across all tumor samples and paired normal tissues.(Dot plot)

Each dots represent expression of samples.



Supplementary Figure 4. The transcription level of Nuf2 in different tumor types through TIMER and GEPIA.





Supplementary Figure 5. Association between CENPL transcription level and immune cell markers. The association between CENPL transcription level and cell markers of (A) CD8+ T cell, (B) B cell, (C) Monocyte, (D–E) T cell, (F) M1 Macrophage, (G) M2 Macrophage, (H–I) DC, (J) TAM, (K–L) NK, (M–N) Neutrophil, (O) Th1, (P) Th2, (Q) Tfh, (R) Th17, (S) Treg, and (T–U) T-cell exhaustion. Abbreviations: DC: dendritic cell; TAM: tumor-associated macrophage; NK: natural killer cell; Th: T helper cell; Tfh: follicular helper T cell; Treg: regulatory T cell.