

Increased Expression Together with the Gene Regulation Network of NEIL3 Predicts Poor Prognosis and Correlates with Immune Infiltrates in Hepatocellular Carcinoma

Bo Hu

Department of Liver Surgery

Xiao-Bo Yang

Department of Liver Surgery

Xinting Sang (✉ sangxt@pumch.cn)

Peking Union Medical College Hospital <https://orcid.org/0000-0003-1952-0527>

Research

Keywords: NEIL3, Hepatocellular Carcinoma, Prognosis, The Cancer Genome Atlas, Immune Infiltrates

Posted Date: February 13th, 2020

DOI: <https://doi.org/10.21203/rs.2.23392/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License. [Read Full License](#)

Abstract

Background: Abnormal Nei endonuclease VIII-like 3 (NEIL3) expression is associated with carcinogenesis. **Methods:** We used sequencing data from the Cancer Genome Atlas database, analyzed NEIL3 expression, gene regulation networks and the correlation with immune infiltrates in hepatocellular carcinoma (HCC). Clinicopathologic characteristics associated with overall survival in TCGA patients using Cox regression and the Kaplan-Meier method. Gene Set Enrichment Analysis was performed using TCGA data set. LinkedOmics was used to identify differential gene expression with NEIL3 and to analyze Gene Ontology and Kyoto Encyclopedia of Genes and Genomes pathways. Gene enrichment analysis examined target networks of kinases and transcription factors. Correlations between NEIL3 expression and cancer immune infiltrates and immune gene markers were analyzed by TIMER and GEPIA. **Results:** We found that overexpressed NEIL3 predicted poor prognosis. Functional network analysis suggested that NEIL3 regulates the DNA replication and cell cycle signaling via pathways involving several cancer-related kinases and E2F Transcription Factor 1. NEIL3 was also found to be associated with the infiltration of several immune cells. **Conclusions:** Our results demonstrate that data mining efficiently reveals information about NEIL3 expression, potential regulatory networks and the relationship with immune infiltration in HCC, laying a foundation for further study of the role of NEIL3 in carcinogenesis.

1. Background

Hepatocellular carcinoma (HCC) is the second leading cause of cancer-related deaths worldwide, which affects about 800,000 cases every year[1]. Currently, HCC can be treated by surgical treatment and chemotherapy, but the mortality rate remains high[2]. The development of various targeted drugs has prolonged patient survival and made a revolutionary breakthrough in treating advanced HCC[3]. Nonetheless, no satisfactory efficacy is attained by the existing targeted drugs. In recent years, immune checkpoint inhibitors (ICIs), such as ipilimumab (the CTLA4 inhibitor) and nivolumab (the PD-1 inhibitor), demonstrate survival benefits for HCC, which reveals the close relationship between immune status and HCC progression[4]. The pathogenesis of HCC is extremely complex, which involves processes like cell cycle regulation and signal transduction, and reflects the function and interaction of multiple genes in multiple steps [5]. Typically, it may be possible to identify the novel drug targets of HCC through screening the gene networks for changes related to tumor genesis and progression.

Nei endonuclease VIII-like 3 (NEIL3), which serves as the DNA glycosylase, is up-regulated at the S phase and peaks at the late S/G2 phase; specifically, it preferentially repairs the oxidative lesions in telomere sequences in vitro [6]. NEIL3 efficiently excises the hydantoin lesions spiroiminodi-hydantoin (Sp) and guanidino-hydantoin (Gh) in single-stranded (ss) and double-stranded (ds) DNA; moreover, it also removes 5-hydroxy-2'-deoxycytidine (5OHC) and 5-hydroxy-2'-deoxyuridine (5OHU) from the ssDNA[7]. Unlike NEIL1 and NEIL2 that possess the β , δ -elimination activity, NEIL3 mainly incises the damaged DNA by means of β -elimination. In addition, the base excision and strand incision activities of NEIL3 were discordant, which indicates that NEIL3 mainly operates as a monofunctional DNA glycosylase. On the other hand, it is proposed that the NEIL3-dependent modulation of DNA methylation can regulate the proliferation of cardiac fibroblasts, thereby affecting lipid metabolism and extracellular matrix (ECM) modulation following

myocardial infarction (MI)[8]. Klattenhoff et al. also demonstrated that, the loss of NEIL3 DNA glycosylase markedly increased the replication-associated double strand breaks (DSBs) and improved the sensitivity to ATR inhibitor in glioblastoma cells[9]. Moreover, aberrant NEIL3 expression is observed in several cancers, and the median NEIL3 expression level is reported to be markedly positively correlated with the median somatic single and total mutation loads, as suggested by data on 13 cancer types from The Cancer Genome Atlas (TCGA) database[10]. The NEIL3 expression level in HCC is higher than that in normal hepatic tissues[11, 12], however, the underlying molecular mechanism, together with its effect on HCC initiation and development, remains largely unclear so far. Moreover, the clinical relevance and prognostic significance of NEIL3 in HCC should be further explored.

In this study, NEIL3 expression and its correlation with the prognosis for HCC patients from databases (such as TCGA, HCCDB, UALCAN and Kaplan-Meier plotter) were comprehensively analyzed. Besides, the functional networks related to NEIL3 in HCC were also examined using the multi-dimensional analysis methods. Furthermore, the correlation of NEIL3 with the tumor-infiltrating immune cells (TIICs) was detected through the Tumor Immune Estimation Resource (TIMER). In a word, our results potentially revealed the novel targets and strategies for the diagnosis and treatment of HCC.

2. Materials And Methods

2.1. Acquisition of clinical samples and patient data for RNA-sequencing

The transcriptomic RNA-sequencing data from HCC samples were downloaded from the TCGA data portal (<https://cancergenome.nih.gov/>), which covered data from 374 primary HCC tissues and 50 non-carcinoma tissues. The NEIL3 expression levels were then extracted. At the same time, the clinicopathological characteristics for these patients, including age, gender, pathological stage (grade), tumor stage, T stage, lymph node metastasis (LNM) and distant metastasis (DM), were also downloaded and collected. In addition, HCCDB (<http://lifeome.net/database/hccdb/about.html>), the one-stop online resource that contains 15 public HCC expression datasets involving up to about 4000 clinical samples for exploring the HCC gene expression, was also adopted for more extensive validation[13]. Supplementary Table 1 presents the correspondence between the HCCDB database and the Gene Expression Omnibus (GEO) database.

2.2. Survival curve plotting and clinical correlation analysis

Clinical data downloaded from TCGA data portal were collected to extract the survival time, and the primary endpoint was deemed as death. Further, the survival curve was plotted using the survival package of R software. Then, the relationship of clinical parameters with NEIL3 expression was also examined to reveal the underlying clinical significance of NEIL3. Noteworthy, the Kaplan-Meier plotter, which can assess the impacts of 54 k genes on the survival for 21 cancer types (<http://kmplot.com/analysis/>), was also utilized to assist in estimating the effect of NEIL3 on survival[14, 15]. The largest datasets in the Kaplan-Meier plotter include breast (n = 6,234), ovarian (n = 2,190), lung (n = 3,452), and gastric (n = 1,440) cancers; whereas the miRNA subsystems cover 11 k samples from 20 cancer types. This system has contained the gene chip and

RNA-seq data sources for databases like GEO, EGA, and TCGA. In this study, the hazard ratio (HR) with the corresponding 95% confidence intervals (CI), and the log-rank P-value were also calculated. In view of the small number of patients at stage 4, patients at stage 3 and 4 were also adopted for analysis. Moreover, UALCAN, the interactive web-portal that uses TCGA level 3 RNA-seq and clinical data from 31 cancer types for in-depth analyses on TCGA gene expression data, was also employed to assist in analysis[16]. Typically, the portal possesses the user-friendly feature of allowing for analysis on the relative expression of a query gene(s) across tumor and normal samples, as well as among various tumor sub-groups according to individual cancer stages, tumor grade or other clinicopathological features. UALCAN is publicly available at <http://ualcan.path.uab.edu>.

2.3. Gene set enrichment analysis (GSEA)

GSEA is a computational approach, which determines whether an a priori defined set of genes shows statistically significant and concordant differences between two biological states[17]. In our study, an ordered list of all genes was generated by GSEA at first based on their correlations with NEIL3 expression; afterwards, GSEA was conducted to illustrate the significant survival difference observed between the high and low NEIL3 expression groups. Typically, the gene set permutations were performed for 1000 times in each analysis, with NEIL3 expression being used as a phenotype label. In addition, the nominal p-value and the normalized enrichment score (NES) were adopted to sort the pathways enriched in each phenotype. Moreover, multiple GSEA was also carried out from a comprehensive perspective.

2.4. LinkedOmics analysis

The LinkedOmics database (<http://www.linkedomics.org/login.php>) is a Web-based platform for analyzing 32 TCGA cancer-associated multi-dimensional datasets[18]. In this study, the LinkFinder module of LinkedOmics was utilized to examine the differentially expressed genes (DEGS) correlated with NEIL3 in TCGA LIHC cohort (n = 371). The Spearman correlation coefficient was used to statistically analyzed the results, and the statistical plots for individual genes were also created. All results were graphically presented in the forms of volcano plots, heat maps or scatter plots. Notably, the LinkInterpreter module of LinkedOmics carries out pathway and network analyses of the DEGs. In this study, data obtained from the LinkFinder results were signed and ranked, and GSEA was employed for Gene Ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway, kinase-target enrichment, and transcription factor-target enrichment analyses. Notably, the final network analyses were conducted based on the Molecular Signatures Database (MSigDB). FDR of < 0.05 was selected as the rank criterion, and 500 simulations were performed.

2.5. TIMER database analysis

TIMER is a web resource to systematically evaluate the clinical impacts of different immunocytes on various cancer types, which covers 10,879 samples across 32 cancer types from TCGA to estimate the abundances of six tumor-infiltrating immunocyte subtypes, including B cells, CD4 T cells, CD8 T cells, macrophages, neutrophils, and dendritic cells (<https://cistrome.shinyapps.io/timer/>)[19]. This database applies the previously published statistical method deconvolution to predict the abundances of tumor-infiltrating immune cells (TIICs) based on the gene expression profiles [20]. This study analyzed the correlation of NEIL3 expression with the abundances of immune infiltrates in HCC, including B cells, CD4 + T

cells, CD8 + T cells, neutrophils, macrophages, and dendritic cells, using the gene modules. Besides, the gene expression levels against tumor purity were also displayed[21]. Moreover, correlations between NEIL3 expression and the gene markers of several TIICs were examined using the correlation modules. Notably, the gene markers of TIICs included markers of CD8 + T cells, T cells (general), B cells, monocytes, TAMs, M1 macrophages, M2 macrophages, neutrophils, T-helper 1 (Th1) cells, T-helper 2 (Th2) cells, follicular helper T (Tfh) cells, T-helper 17 (Th17) cells, Tregs, and exhausted T cells. These gene markers had been reported in previous studies [22–26]. NEIL3 expression was used as the x-axis with gene symbols, whereas related marker genes were displayed on the y-axis as gene symbols. The gene expression level was expressed in the form of log₂ RSEM.

2.6. Gene correlation analysis in GEPIA

The online database Gene Expression Profiling Interactive Analysis (GEPIA) (<http://gepia.cancer-pku.cn/index.html>) is an interactive web, which covers 9,736 tumors and 8,587 normal samples from TCGA and the Genotype-Tissue Expression (GTEx) projects, and analyses the RNA sequencing data[27]. In this study, GEPIA was adopted to further confirm genes with significant correlation in TIMER. Additionally, gene expression correlation analysis was carried out on the given sets of TCGA expression profile data. The correlation coefficient was determined according to the Spearman method. NEIL3 was used as the x-axis, while the other genes of interest were presented on the y-axis, and both tumor and normal tissue datasets were used for analysis.

2.7. Statistical analysis

The R (v.3.6.1) package was adopted for all statistical analyses. The relationships of clinicopathological features with NEIL3 expression were examined through the Wilcoxon signed-rank test and logistic regression. Thereafter, the clinicopathological characteristics associated with the OS for TCGA patients were determined using Cox regression and the Kaplan-Meier method. Further, the effects of NEIL3 expression on patient survival along with other clinical characteristics (such as age, gender, stage, grade, T stage, LNM status, and DM status) were compared using the multivariate Cox analysis. The threshold NEIL3 expression was determined by its median level. Besides, the results of Kaplan-Meier plots, UALCAN and GEPIA were presented as HRs and P-values or Cox P-values upon the log-rank test. Moreover, Spearman's correlation and statistical significance was conducted to evaluate the correlations of gene expression. P-values of < 0.05 were considered as statistically significant.

3. Result

3.1. NEIL3 mRNA expression in HCC and other human cancers

To determine the differences in NEIL3 expression between HCC and normal tissues, the differential scatter plot and paired plot were obtained (Fig. 1A and 1B). Also, NEIL3 expression among different GEO datasets was analyzed using the HCCDB database. Furthermore, NEIL3 mRNA expression within various cancer types and matched normal tissues were measured based on the TIMER database, so as to reveal the aberrant

NEIL3 expression from a holistic perspective. Our analysis suggested that NEIL3 expression was up-regulated in HCC among different datasets (Fig. 1C), as well as in bladder, breast, biliary, colorectal, esophageal, gastric, kidney, lung, prostatic and rectal cancers compared with that in normal tissues (Fig. 1D).

3.2. Survival outcomes and multivariate analysis

The Kaplan-Meier survival analysis in Fig. 2A suggested that HCC with high NEIL3 expression was associated with worse prognosis than that with low NEIL3 expression in terms of TCGA ($p < 0.001$). Similar results were also obtained from the analysis of HCC samples in the International Cancer Genome Consortium (ICGC) HCC cohort (Fig. 2B). The values of area under the receiver operating characteristic (ROC) curve (AUC) for 1-, 3- and 5-year OS of TCGA patients were 0.596, 0.663 and 0.737, respectively (Fig. 2C). In addition, to better illustrate the relevance and underlying mechanisms of NEIL3 expression in HCC, the relationships of NEIL3 expression with clinical characteristics and survival prognosis for HCC patients were analyzed in the Kaplan-Meier plotter database (Table 1). According to our results, NEIL3 over-expression was related to worse OS in male and female patients, white and Asian patients, together with those with different clinical stages and T stages upon the American Joint Committee on Cancer (AJCC), none or micro-vascular invasion status, sorafenib treatment as well as both risk factors ($P < 0.05$). Specifically, the effects of high NEIL3 expression on OS were particularly significant in Asian (HR = 6.54, $P < 0.001$) and G1 (HR = 4.18, $P < 0.01$) patients.

Clinical characteristics		Total(N)	HR (95%CI)	p-Value
Grade	G2 vs. G1	235	2.31 (1.20 - 4.66)	*
	G3 vs. G1	178	6.47 (3.23 - 13.64)	***
	G4 vs. G1	67	4.10 (1.13 - 15.93)	*
Stage	II vs. I	262	2.15 (1.28 - 3.68)	**
	III vs. I	257	1.92 (1.14 - 3.26)	*
	IV vs. I	180	0.35 (0.018 - 2.44)	0.355
T stage	T2 vs. T1	280	2.55 (1.54 - 4.30)	***
	T3 vs. T1	265	1.85 (1.09 - 3.16)	*
	T4 vs. T1	197	3.40 (1.07 -12.95)	*
Distant metastasis	positive vs. negative	276	0.33 (0.02 - 2.60)	0.338
Lymph nodes	positive vs. negative	261	0.33 (0.02 - 2.60)	0.337

HR, Hazard ratio; CI, confidence interval.

***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$.

Table 1. NEIL3 expression associated with clinical pathological characteristics (logistics regression)

On the other hand, univariate analysis revealed that high NEIL3 expression was dramatically correlated with the poor OS (HR: 1.441; 95% confidence interval [CI]: 1.249 – 1.664; $p < 0.001$). Besides, clinical stage and T stage were also associated with the poor patient survival (Fig. 4A). Multivariate analysis demonstrated that, NEIL3 was the only factor independently associated with OS, with the HR of 1.357 (CI: 1.162 – 1.585, $p < 0.001$) (Fig. 4B).

3.3. Associations between NEIL3 expression and clinicopathological variables

Altogether 371 HCC samples with available NEIL3 expression data stratified based on multiple patient characteristics were analyzed from TCGA. Figure 3A consistently suggested the high transcription of NEIL3. As shown in Fig. 3B–3G, the transcription levels of NEIL3 were notably higher in HCC patients than in healthy people upon subgroup analyses stratified according to gender, age, weight and obese, ethnicity, disease stage, tumor grade and nodal metastasis status. Additionally, the univariate analysis using logistic regression revealed that NEIL3 expression, which served as a categorical dependent variable (based on median expression level of 2.5), was linked with the clinicopathological characteristics predicting dismal prognosis (Table 2). Moreover, the up-regulated NEIL3 expression in HCC was dramatically associated with high grade (OR = 2.31 for G2 vs. G1, 6.47 for G3 vs. G1, 4.10 for G4 vs. G1), as well as advanced clinical stage (OR = 2.15 for IV vs. I, 1.92 for III vs. I), and T stage (OR = 2.55 for T2 vs. T1, 1.85 for T3 vs. T1, 3.40 for T4 vs. T1). The above findings demonstrated that HCC with high NEIL3 expression was more likely to progress to a more advanced and malignant stage than HCC with low NEIL3 expression.

clinicopathological characteristics	Overall survival		
	(n=364)		
	N	HR (95%CI)	p-Value
SEX			
female	121	2.15 [1.18-3.95]	*
male	250	2.83 [1.71-4.68]	***
RACE			
white	184	1.72 [1.09-2.73]	*
Asian	156	6.54 [2.57-16.6]	***
STAGE			
I	171	2.18 [1.15-4.12]	*
II	86	3.48 [1.04-11.64]	*
III	85	4.51 [2.34-8.72]	***
III+IV	90	4.17 [2.21-7.88]	***
AJCC T			
T1	181	2.01 [1.1-3.66]	*
T2	94	2.85 [1.16-7.03]	**
T3	80	3.45 [1.85-6.44]	***
GRADE			
G1	55	4.18 [1.5-11.68]	**
G2	177	2.1 [1.24-3.57]	**
G3	122	3.51 [1.9-6.47]	***
VASCULAR INVASION			
none	205	2.36 [1.37-4.06]	**
micro	93	2.36 [0.99-5.62]	*
RISK FACTORS			
alcohol consumption	117	2.07 [1.08-3.96]	*
hepatitis virus	153	3.09 [1.52-6.25]	***
TREATMENT			
Sorafenib treatment	30	14.61 [1.76-121.04]	**

HR, hazard ratio; CI, confidence interval.

***, $P < 0.001$; **, $P < 0.01$; *, $P < 0.05$.

Table 2. Correlation of *NEIL3* mRNA expression and clinical prognosis in liver cancer with different clinicopathological factors by Kaplan-Meier plotter.

3.4. GSEA identified the NEIL3-related signaling pathways

To identify the NEIL3-related signaling pathways that were differentially activated in HCC, GSEA was conducted between the low and high NEIL3 expression data sets. The results of GSEA revealed significant differences ($FDR < 0.05$, $NOM\ p < 0.05$) in the enrichment of MSigDB Collection (c2.cp.biocarta and h. all. v6.1. symbols). Results of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis revealed that the genes in high NEIL3 expression group were mainly enriched in the cell cycle, base excision repair and purine metabolism ($FDR = 0.01$, $NOM = 0$) (Fig. 5A-5C); in addition, the complement and coagulation cascades, primary bile acid biosynthesis, valine leucine and isoleucine degradation were primarily enriched in low ANLN expression group ($FDR < 0.05$, $NOM < 0.005$) (Fig. 5D-5F). The markedly enriched signaling pathways related to cancer and immunity, including the P53 pathway, the T cell receptor pathway, the mTOR pathway, the ERBB pathway, the Notch signaling pathways and Pathways in cancer were also differentially enriched in the high NEIL3 expression phenotype (Supplementary Table 2). Afterwards, multiple GSEA analysis was also conducted to further demonstrate the whole picture of the signaling pathways related to NEIL3 expression in HCC (Fig. 5H).

3.5. GO and KEGG pathway analyses of the co-expressed genes correlated with NEIL3 in HCC

The Function module of LinkedOmics was adopted to analyze the mRNA sequencing data from 371 LIHC patients in TCGA. As shown in the volcano plot (Fig. 6A), 129,55 genes (denoted by the dark red dots) were significantly positively correlated with NEIL3, whereas 6,965 (denoted by the dark green dots) displayed markedly negative correlations (false discovery rate [FDR] of < 0.01). Of them, the top 50 significant gene sets exhibiting positive and negative correlations with NEIL3 were presented in the heat map (Fig. 6B and 6C). Such results suggested that NEIL3 had widespread impact on the transcriptome. Supplementary Fig. 1A-1C shows the statistical scatter plots for individual genes. As was figured out, NEIL3 expression was strongly positively correlated with expression of the Centrosome-associated protein E (CENPE, which coupled the chromosome position to microtubule depolymerizing activity) (positive rank #1, Spearman's correlation = 0.774, $p = 2.512e-75$), the non-structural maintenance of chromosomes condensin I complex subunit G (NCAPG, which was responsible for the condensation and stabilization of chromosomes during mitosis and meiosis) (Spearman's correlation = 0.773, $p = 3.568e-75$), and the kinesin family member 4A (KIF4A, which is involved in the intracellular transport of membranous organelles) (Spearman's correlation = 0.773, $p = 6.256e-75$). Additionally, the significant GO term analysis upon GSEA demonstrated that, the DEGs correlated with NEIL3 were mainly located in the condensed chromosome, chromosomal region and

spindle, where they played primary roles in chromosome segregation, spindle organization, cell cycle checkpoint control and DNA replication. Besides, the catalytic and helicase activities, together with single-stranded DNA binding, were involved in the molecular functions (Fig. 7A–7C). KEGG pathway enrichment analyses illustrated the enrichments in cell cycle, DNA replication and homologous recombination pathways. Furthermore, the P53 signaling, base excision repair and the hepatitis pathways were also mentioned (Fig. 7D).

3.6. The NEIL3 networks of kinase and transcription factor targets in HCC

To further explore the targets of NEIL3 in HCC, the kinase and transcription factor target networks were analyzed for the positively correlated gene sets generated by GSEA. Typically, the top 5 most significant target networks were the kinase-target networks primarily related to the polo like kinase 1 (PLK), cyclin-dependent kinase 1 (CDK1), aurora kinase B (AURKB), checkpoint kinase 1 (CHEK1) and aurora kinase A (AURKA) (Table 3). In addition, the transcription factor-target network was mainly related to the E2F Transcription Factor (E2F) family, including E2F1_Q6, E2F1_Q4 and E2F4DP1_01. The protein-protein interaction (PPI) network constructed by GeneMANIA revealed correlations among genes for kinases PLK and TF E2F1_Q6; besides, the gene sets enriched for them were mainly responsible for nuclear division, organelle fission, mitosis and DNA replication. (Fig. 8 and Supplementary Fig. 2).

Table 3

The Kinase, miRNA and transcription factor-target networks of NEIL3 in hepatocellular carcinoma (LinkedOmics)

Enriched category	Gene set	leadingEdgeNum	FDR
Kinase Target	Kinase_PLK1	33	0
	Kinase_CDK1	84	0
	Kinase_AURKB	26	0
	Kinase_CHEK1	41	0
	Kinase_AURKA	16	0
Transcription factor target	V\$E2F1_Q6	82	0
	V\$E2F1_Q4	80	0
	V\$E2F4DP1_01	81	0
	V\$E2F1DP1_01	89	0
	V\$E2F4DP2_01	89	0
LeadingEdgeNum, the number of leading-edge genes; FDR, false discovery rate from Benjamini and Hochberg from gene set enrichment analysis (GSEA); V\$, the annotation found in Molecular Signatures Database (MSigDB) for transcription factors (TF).			

3.7 NEIL3 expression was correlated with the immune infiltration level in HCC

GSEA analysis showed that, the T cell receptor pathway was closely related to NEIL3; as a result, the relationships between NEIL3 expression and the immune infiltration levels in HCC were also examined. Meanwhile, the correlations of NEIL3 expression with the immune infiltration levels in HCC were also examined from TIMER, and the results suggested that NEIL3 expression was significantly correlated with the infiltration levels of B cells ($r = 0.37$, $p = 1.26e^{-12}$), dendritic cells ($r = 0.344$, $p = 7.17e^{-11}$), CD8 + T cells ($r = 0.232$, $p = 1.14e^{-05}$), CD4 + T cells ($r = 0.141$, $p = 8.69e^{-03}$), macrophages ($r = 0.296$, $p = 2.46e^{-08}$) and neutrophils ($r = 0.273$, $p = 2.71e^{-07}$) (Fig. 9). Moreover, functional annotation was further carried out by means of GSEA, and the results suggested that the differentially expressed genes between these two groups were enriched in the immunological signature gene sets (c7. All. V7.0. symbol) (Supplementary Table 3). Taken together, these findings suggested that NEIL3 might play a specific role in the immune infiltration in HCC, especially for B cells, T cells and dendritic cells.

3.8. Correlation analysis between NEIL3 expression and immune marker sets

To further probe into the relationships between NEIL3 and the diverse immune infiltrating cells, the correlations of NEIL3 with the immune marker sets of various HCC immunocytes were explored in the TIMER and GEPIA databases. As suggested by our results, NEIL3 expression was correlated with the immune marker genes of various immune cells, including CD8 + T cells, T cells (general), B cells, monocytes, TAMs, M1 and M2 macrophages, and neutrophils in HCC (Table 4). Also, the diverse functional T cells were also analyzed, such as Th1 cells, Th2 cells, Tfh cells, Th17 cells, Tregs, and the exhausted T cells. Results of correlation analysis adjusted by purity revealed that, the NEIL3 expression level was significantly correlated with the most immune marker sets of various immune cells and the diverse functional T cells in HCC. Moreover, those markers associated with T cells were selected to analyze their correlations with NEIL3 expression in the GEPIA database. According to our results, NEIL3 expression was markedly correlated with the markers related to T cells in HCC, but not with most of them in normal tissues (Table 5), indicating that NEIL3 might be involved in regulating T cell function.

Table 4

The relationship between NEIL3 and relate genes and markers of immune cells in TIMER.

Description	Gene marker	LIHC				CHOL			
		None		Purity		None		Purity	
		Cor	p-Value	Cor	p	Cor	p	cor	p-Value
CD8 + T cell	CD8A	0.095	0.0667	0.116	*	0.114	0.508	-0.019	0.913
	CD8B	0.108	*	0.133	*	0.212	0.213	0.128	0.464
T cell(general)	CD3D	0.201	***	0.235	***	-0.118	0.493	-0.322	0.059
	CD3E	0.118	*	0.159	**	0.039	0.819	-0.168	0.336
	CD2	0.123	*	0.163	**	0.062	0.720	-0.13	0.455
B cell	CD19	0.198	***	0.218	***	0.049	0.776	-0.108	0.526
	CD79A	0.067	0.195	0.082	0.128	0.140	0.414	-0.003	0.988
Monocyte	CD86	0.183	***	0.256	***	0.020	0.906	-0.156	0.372
	CD115	0.0069	0.183	0.125	*	0.011	0.950	-0.124	0.477
	CD14	-0.225	***	-0.233	***	0.059	0.731	-0.108	0.535
TAM	CCL2	-0.097	0.0632	-0.095	0.0781	-0.138	0.421	-0.215	0.215
	CD68	0.156	*	0.187	***	0.149	0.383	0.076	0.664
	IL10	0.121	*	0.167	***	0.337	0.044	0.208	0.231
M1 Macrophage	NOS2	-0.163	*	-0.167	***	0.007	0.970	0.009	0.96
	IRF5	0.123	*	0.11	*	-0.038	0.825	-0.121	0.489
	CD11c	-0.105	***	0.284	***	0.021	0.904	-0.16	0.357
	CD80	0.302	***	0.356	***	0.159	0.353	0.023	0.896
M2 Macrophage	CD206	-0.152	**	-0.14	**	0.155	0.364	0.039	0.823
	VSIG4	0.054	0.299	0.0096	0.0743	0.154	0.369	0.035	0.842

LIHC, liver hepatocellular carcinoma; CHOL, cholangiocarcinoma; TAM, tumor-associated macrophage; Tfh, Follicular helper T cell; Treg, regulatory T cell; Cor, R value of Spearman's correlation; None, correlation without adjustment. Purity, correlation adjusted by purity.

***, P < 0.001; **, P < 0.01; *, P < 0.05.

Description	Gene marker	LIHC				CHOL			
		None		Purity		None		Purity	
		Cor	p-Value	Cor	p	Cor	p	cor	p-Value
	MS4A4A	0.053	0.313	0.0094	0.0799	0.151	0.378	-0.05	0.774
	CD200R1	0.165	**	0.202	***	0.022	0.897	-0.16	0.359
	FCGR2A	0.222	***	0.25	***	0.057	0.740	-0.028	0.872
	FCGR2B	0.044	0.4	0.051	*	0.055	0.748	-0.138	0.431
Neutrophils	CD66b	0.043	0.414	0.048	0.378	0.215	0.208	0.218	0.208
	CD11b	0.303	***	0.048	***	0.034	0.844	-0.041	0.813
	CD15	0.115	*	0.121	*	-0.071	0.682	-0.104	0.552
	CD13	-0.225	***	-0.233	***	-0.155	0.364	-0.133	0.448
	CCR7	-0.015	0.773	0.014	0.801	0.076	0.657	-0.116	0.507
Th1	STAT4	0.129	*	0.141	**	0.006	0.972	-0.082	0.64
	STAT1	0.135	**	0.14	**	-0.030	0.863	-0.101	0.565
	IFNG	0.225	***	0.257	***	0.176	0.304	0.05	0.776
	TNF	0.183	***	0.231	***	-0.103	0.547	-0.176	0.312
	CCR5	0.159	**	0.218	***	0.075	0.663	-0.132	0.45
Th2	CD30	0.203	***	0.262	***	0.058	0.736	-0.098	0.575
	GATA3	0.059	0.259	0.105	0.0518	0.001	0.998	-0.191	0.271
	STAT6	-0.211	***	-0.202	***	-0.240	0.158	-0.263	0.126
	IL1RL1	-0.294	***	-0.319	***	0.062	0.720	-0.035	0.841
	CCR3	0.33	***	0.366	***	0.175	0.308	0.08	0.647
	CCR4	0.103	*	0.123	*	0.183	0.285	0.051	0.771
Tfh	IL21	0.144	**	0.155	**	-0.065	0.704	-0.157	0.366
	CXCL13	0.177	***	0.205	***	0.206	0.229	0.096	0.582

LIHC, liver hepatocellular carcinoma; CHOL, cholangiocarcinoma; TAM, tumor-associated macrophage; Tfh, Follicular helper T cell; Treg, regulatory T cell; Cor, R value of Spearman's correlation; None, correlation without adjustment. Purity, correlation adjusted by purity.

***, P < 0.001; **, P < 0.01; *, P < 0.05.

Description	Gene marker	LIHC				CHOL			
		None		Purity		None		Purity	
		Cor	p-Value	Cor	p	Cor	p	cor	p-Value
	ICOS	0.266	***	0.323	***	0.126	0.465	-0.04	0.822
Treg	FOXP3	0.063	0.224	0.09	0.0942	0.146	0.395	-0.006	0.973
	CCR8	0.353	***	0.392	***	0.123	0.476	-0.001	0.996
	STAT5B	-0.123	*	-0.142	**	0.188	0.271	0.144	0.41
T cell exhaustion	PDCD1	0.216	***	0.236	***	0.117	0.496	0.042	0.813
	PD-L1	0.122	*	0.146	**	0.114	0.506	-0.019	0.912
	CTLA4	0.303	***	0.356	***	0.080	0.641	-0.018	0.917
	LAG3	0.145	**	0.154	**	0.268	0.114	0.191	0.273
	TIM-3	0.233	***	0.314	***	0.058	0.734	-0.086	0.624
LIHC, liver hepatocellular carcinoma; CHOL, cholangiocarcinoma; TAM, tumor-associated macrophage; Tfh, Follicular helper T cell; Treg, regulatory T cell; Cor, R value of Spearman's correlation; None, correlation without adjustment. Purity, correlation adjusted by purity.									
***, P < 0.001; **, P < 0.01; *, P < 0.05.									

Table 5

Correlation analysis between NEIL3 and relate genes and markers of T cells in GEPIA.

Description	Gene marker	LIHC				CHOL			
		Tumor		Normal		Tumor		Normal	
		R	p-Value	R	p-Value	R	p-Value	R	p-Value
CD8 + T cell	CD8A	0.18	***	0.26	0.063	0.18	0.28	0.52	0.25
	CD8B	0.15	**	0.23	0.1	0.25	0.14	0.27	0.49
T cell(general)	CD3D	0.25	***	0.26	0.063	0.024	0.89	0.25	0.52
	CD3E	0.2	***	0.25	0.076	0.072	0.68	0.29	0.44
	CD2	0.2	***	0.27	0.057	0.13	0.74	0.42	0.26
TCR	CD28	0.32	***	0.15	0.21	0.14	0.42	0.6	0.086
	CD45	0.32	***	0.27	0.057	0.18	0.31	0.46	0.21
	CD4	0.24	***	0.11	0.47	0.11	0.54	0.27	0.47
Th1	STAT4	0.24	***	0.18	0.21	0.02	0.91	0.65	0.059
	STAT1	0.31	***	0.21	0.14	0.11	0.52	-0.044	0.91
	IFNG	0.28	***	0.11	0.45	0.17	0.32	0.14	0.71
	TNF	0.29	***	0.27	0.061	-0.057	0.74	0.14	0.71
	CCR5	0.27	***	0.29	*	0.15	0.37	0.42	0.26
Th2	CD30	0.29	***	0.19	0.18	0.17	0.33	0.49	0.18
	GATA3	0.19	***	0.094	0.52	0.025	0.89	0.5	0.17
	STAT6	0.15	**	0.13	0.37	0.011	0.95	0.36	0.34
	IL1RL1	-0.19	***	0.045	0.76	0.048	0.78	-0.4	0.29
	CCR3	0.37	***	-0.13	0.37	0.28	0.096	0.66	0.054
	CCR4	0.23	***	0.17	0.23	0.2	0.25	-0.035	0.93
Tfh	IL21	0.17	***	0.14	0.32	-0.001	0.99	0.073	0.85
	CXCL13	0.22	***	0.049	0.74	0.24	0.15	-0.11	0.79

LIHC, liver hepatocellular carcinoma; CHOL, cholangiocarcinoma; TAM, tumor-associated macrophage; Tfh, Follicular helper T cell; Treg, regulatory T cell; Tumor, correlation analysis in tumor tissue of TCGA. Normal, correlation analysis in normal tissue of TCGA.

***, P < 0.001; **, P < 0.01; *, P < 0.05.

Description	Gene marker	LIHC				CHOL			
		Tumor		Normal		Tumor		Normal	
		R	p-Value	R	p-Value	R	p-Value	R	p-Value
	ICOS	0.33	***	0.22	0.12	0.2	0.24	0.2	0.61
Treg	FOXP3	0.13	*	0.4	**	0.2	0.24	0.38	0.31
	CCR8	0.44	***	0.15	0.29	0.16	0.35	0.15	0.7
	STAT5B	0.25	***	0.021	0.89	0.25	0.14	-0.18	0.65
T cell exhaustion	PDCD1	0.28	***	0.33	*	0.26	0.13	0.31	0.42
	PD-L1	0.23	***	0.18	0.22	0.21	0.21	0.089	0.82
	CTLA4	0.37	***	0.32	*	0.23	0.17	0.21	0.59
	LAG3	0.18	***	0.18	0.21	0.3	0.076	0.37	0.32
	TIM-3	0.36	***	0.12	0.42	0.22	0.2	0.2	0.62
LIHC, liver hepatocellular carcinoma; CHOL, cholangiocarcinoma; TAM, tumor-associated macrophage; Tfh, Follicular helper T cell; Treg, regulatory T cell; Tumor, correlation analysis in tumor tissue of TCGA. Normal, correlation analysis in normal tissue of TCGA.									
***, P < 0.001; **, P < 0.01; *, P < 0.05.									

4. Discussion

The genetic abnormalities of NEIL3, a DNA glycosylase that initiates base excision repair by hydrolyzing the N-glycosidic bond and releasing the damaged base, are reported to be correlated with hepatocarcinogenesis[28]. Recent study also demonstrates that the Neil3 gene expression levels are highly related to HCC recurrence in patients receiving hepatectomy, which suggests that NEIL3 may serve as a prognostic marker for HCC [11]. To the best of our knowledge, NEIL3 expression, together with its potential prognostic impact on HCC, has not yet been explored, and the potential role of NEIL3 in HCC is the focus in this study.

As found in this study, NEIL3 was highly expressed in HCC, as well as bladder, breast, biliary, head and neck, colorectal, esophageal, gastric, kidney, lung, uterus, prostatic, and rectal cancers, compared with that in normal tissues. The differential NEIL3 expression was further observed in numerous independent datasets with regard to HCC based on the HCCDB database. Emphatically, bioinformatic analysis using the high throughput RNA-sequencing data from TCGA demonstrated that, the up-regulated NEIL3 expression in HCC was related to the advanced clinicopathological characteristics (such as high grade, clinical stage and T stage), short survival time, and dismal prognosis, which were supported by the results acquired from UALCAN. Furthermore, the Kaplan-Meier Plotter data suggested that high NEIL3 expression was correlated

with a high hazard ratio (HR) for poor OS and PFS of HCC. To be specific, the up-regulated NEIL3 expression was correlated with the poor prognosis for HCC at stage 2–4, grade 1–3 and T1-T3. Moreover, patients with micro-vascular invasion, alcohol consumption, hepatitis virus infection and Sorafenib treatment also had worse OS and PFS in the presence of high NEIL3 expression. In addition, as far as HR was concerned, the prognosis for male and Asian patients with up-regulated NEIL3 expression was worse than that for female and white patients. Taken together, these findings strongly suggested that NEIL3 was a prognostic biomarker for HCC.

To further investigate the functions of NEIL3 in HCC, GSEA was implemented using TCGA data, and then the differentially enriched pathways in the high NEIL3 expression phenotype as well as those linked to cancer and immunity, were selected. Among them, Base excision repair scored 2.17 on NES, and the polymorphisms of base-excision repair genes are reported to shape the development of hepatitis C virus (HCV) infection, liver cirrhosis and HCC, which predicts the OS[29, 30]. Besides, plenty of studies have elaborated that, the Notch signaling that may induce epithelial-mesenchymal transition (EMT) and mutation or loss of P53, the ErbB pathway that regulates the fundamental cellular processes, and the mTOR signaling that determines cell fate at different levels, play crucial roles in hepatocarcinogenesis[31–33]. Nevertheless, this study was the first to propose that, NEIL3 might be involved in these pathways in the HCC setting, suggesting that NEIL3 might serve as a potential prognostic marker and therapeutic target of HCC.

To further probe the correlations between NEIL3 and other genes in HCC, the interactions between NEIL3 and other functional partners were investigated in the LinkedOmics database. Our results revealed that, there were plenty of genes co-expressed with NEIL3 in LIHC, among which, CENPE, a spindle checkpoint protein, exhibited the highest Spearman's correlation of 0.774. Of interest, CENPE has been identified as a tumor suppressor gene, and its decreased expression may contribute to HCC development[34]. However, recent study documents that CENPE is remarkably up-regulated in HCC[35], consistent with our results on CENPE in TCGA. Therefore, CENPE was speculated to play a dual role in HCC progression. NCAPG, which ranks the second in our list, has shown its effect on HCC cell proliferation and migration, and it is evidently correlated with HCC recurrence, metastasis, differentiation and TNM stage[36]. Additionally, Hu et al. demonstrated that the over-expression of KIF4A, which was also highly correlated with NEIL3, was regulated by forkhead box M1 (FOXM1); besides, it had a bearing on HCC recurrence and progression, and might serve as a promising prognostic marker[37]. Furthermore, enrichment analyses on target gene sets using GSEA helps to reveal the important networks of target kinases and transcription factors (TFs). Our results suggested that the functional network of NEIL3 mainly participated in chromosome segregation, spindle organization, cell cycle regulation, and DNA replication. Interestingly, NEIL3 is also proposed in previous study to be involved in the cell signaling pathway that senses oxidative stress (OS), and contributes to the recruitment of TFs as well as the introduction of chromatin modifications, thereby resulting in the fine tuning of cellular processes for promoting cell survival and successful cell differentiation[38]. Moreover, among the highly NEIL3-related kinases investigated in this study, PLK1, one of the conserved serine/threonine kinase family members involved in multiple mitotic processes (including functional maturation of centrosomes, establishment of the bipolar spindle, chromosome segregation and response to DNA damage), was suggested in previous study to be a factor that determined the SN38 sensitivity of tumor cells, and its over-expression was inversely correlated with the survival rate of HCC[39]. Besides, other kinases, such as CDK1 and AURKB,

which are involved in our network, also regulate genomic stability, mitosis and the cell cycle[40, 41]. Among the TFs targeted by NEIL3, E2F1 represents a key link in the cell cycle regulation network, and its aberrant expression participates in HCC occurrence and development, which also predicts the unfavorable patient prognosis[42]. As reported, E2F1, which is tightly linked to cell division cycle associated 5 (CDCA5), Sirtuin 5 (SIRT5) and other important molecules, may serve as a vital anti-apoptotic factor in liver cancer due to its ability to offset c-Myc-driven apoptosis[43, 44]. As suggested in previous literature, the repressive complex, which includes the component of E2F4, is relieved by CDK upon re-entry into the cell cycle; thereafter, E2F1-Sp1, which is in the form of complex E2F1-NF- κ B, can be recruited to the promoter[45]. Therefore, it was deduced in this study that NEIL3 might regulate this process. Therefore, our analyses suggested that PLK, CDK, E2F1, and E2E4 were all the concerned targets of NEIL3, and that NEIL3 acted through these factors to regulate the cell cycle and proliferation capacity of HCC. Nonetheless, further studies are warranted to test this hypothesis.

Another important discovery of this study was that, NEIL3 expression was correlated with the diverse immune infiltration levels in HCC. As demonstrated by our results, the infiltration levels of macrophages, DCs, B cells, CD8 + T cells, CD4 + T cells and neutrophils showed significantly positive correlations with NEIL3 expression in HCC. Moreover, the correlations of NEIL3 expression with the marker genes of immune cells implicated that NEIL3 might be involved in some immunocyte functions in HCC. In this study, cholangiocarcinoma (CHOL), in which NEIL3 was not associated with immune cell infiltration, was applied as the control. In general, after being adjusted by purity, marker genes related to T cells such as CCR8, CCR3, CTLA4, ICOS and TIM-3, had higher correlations with NEIL3 expression than those of other immunocyte types, while no significant correlation was observed in the control group. What's more, results obtained from GEPIA database analysis indicated that, NEIL3 might potentially affect T cell exhaustion, as well as T cell receptor (TCR) and Treg functions, as supported by the evidence that the up-regulated NEIL3 expression was positively correlated with the expression of most Treg, TCR and T cell exhaustion markers (like SHNP2, CCR8, TIM-3, CTLA4, CD28, CD45 and PD-1). PD-1, a negative costimulatory receptor that is critical for suppressing T cell activation both in vitro and in vivo, becomes clustered with TCR upon binding to its ligand PD-L1, which is transiently associated with the phosphatase SHP2[46]. Typically, NEIL3 may exert its own strength to participate in this system. CCR8, together with CCR10, is the chemokine receptor responsible for Treg cell migration to the tumor microenvironment (TME) [47], and high NEIL3 expression is also proved to be linked with this molecule. On the other hand, TIM-3, a critical surface protein on the exhausted T cells[48], shows positive correlation with NEIL3 expression in HCC. On the other hand, NEIL3 expression is also markedly correlated with the regulation of several T helper cells (Th1, Th2, Tfh, and Th17) markers in HCC. These correlations may indicate a potential mechanism by which NEIL3 regulates T cell functions in HCC. An accumulating amount of data indicate that, DNA glycosylases in general are involved in the adaptive immunity. However, few previous studies are carried out to examine the relationship of NEIL3 with immunity. For instance, Torisu et al. had proposed a putative role of NEIL3 in lymphocytes and/or other immune cells[49]. Meanwhile, other study demonstrates that Neil3-deficient mice are linked with signs of inflammation and autoimmunity when they are challenged by inflammatory stimuli, and that individuals with NEIL3 mutation show certain subclinical signs of autoimmunity[50]. Nevertheless, it remains to be verified about the role of NEIL3 in humans with immunodeficiency and autoimmune disorders. In a word,

although the relationship of NEIL3 over-expression with immune cell infiltration is generally not strong, this study has opened up a new field of vision regarding the influence of NEIL3 on HCC and its underlying role in recruiting and regulating the immune infiltrating cells in HCC.

Nonetheless, some limitations should be noted in this study, which should be taken into consideration when interpreting our results. Firstly, transcriptomic analysis only reflected some aspects of genes, rather than the global alterations. Secondly, there were few stage 4 patients in the LIHC samples, since most HCC patients were first diagnosed at the advanced stage with dismal prognosis. Finally, our results were not validated via another independent cohort, which was also a limitation in this study, and the reliability of our molecular results was still challenged due to the lack of experiments in vitro or in vivo.

In conclusion, NEIL3, which may be positively co-expressed with CEPNE, NCAPG and KIF14, can serve as a potential prognostic molecular marker for the poor survival of HCC. Moreover, the base excision repair, the P53 pathway, the T cell receptor pathway, the mTOR pathway, the ERBB pathway and the Notch signaling may be the key pathways that are regulated by NEIL3; whereas PLK, CDK, E2F1_Q6 and E2F1_Q4 in EC may be targeted by NEIL3, and are involved in cell cycle regulation and DNA replication. Additionally, NEIL3 is also found to be associated with the infiltration of immune cells, especially for the functions of certain T cells related to HCC. Further experimental validation should be performed to confirm the biological impact of NEIL3.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

All data generated or analyzed during this study are included in this article and its supplementary information files.

Competing interests

The authors declare that they have no competing interests.

Funding

Not applicable.

Authors' contributions

Xin-Ting Sang and Bo Hu created the idea for the review. Bo Hu performed the selection of literature, drafted the manuscript, and prepared the figures. Xiao-Bo Yang and Xin-Ting Sang revised the manuscript. All authors read and approved the final manuscript.

Acknowledgments

The authors would like to thank the HCCDB, UALCAN, Kaplan-Meier plotter, TIMER, LinkedOmics and TCGA databases for the availability of the data.

References

1. Ferlay, J., et al., Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *International journal of cancer*, 2015. 136(5): p. E359-E386.
2. Llovet, J.M., R. Montal, and A. Villanueva, *Randomized trials and endpoints in advanced HCC: Role of PFS as a surrogate of survival*. *Journal of hepatology*, 2019.
3. MA, A., et al., *Dual-targeted casein micelles as green nanomedicine for synergistic phytotherapy of hepatocellular carcinoma*. *Journal of controlled release : official journal of the Controlled Release Society*, 2018. **287**(undefined): p. 78-93.
4. B, H., Y. XB, and S. XT, *Liver graft rejection following immune checkpoint inhibitors treatment: a review*. *Medical oncology (Northwood, London, England)*, 2019. **36**(11): p. 94.
5. Levrero, M. and J. Zucman-Rossi, *Mechanisms of HBV-induced hepatocellular carcinoma*. *Journal of hepatology*, 2016. **64**(1): p. S84-S101.
6. J, Z., et al., *NEIL3 Repairs Telomere Damage during S Phase to Secure Chromosome Segregation at Mitosis*. *Cell reports*, 2017. **20**(9): p. 2044-2056.
7. SZ, K., et al., *Human NEIL3 is mainly a monofunctional DNA glycosylase removing spiroimidiohydantoin and guanidinohydantoin*. *DNA repair*, 2013. **12**(12): p. 1159-64.
8. MB, O., et al., *NEIL3-Dependent Regulation of Cardiac Fibroblast Proliferation Prevents Myocardial Rupture*. *Cell reports*, 2017. **18**(1): p. 82-92.
9. AW, K., et al., *Loss of NEIL3 DNA glycosylase markedly increases replication associated double strand breaks and enhances sensitivity to ATR inhibitor in glioblastoma cells*. *Oncotarget*, 2017. **8**(68): p. 112942-112958.
10. K, S., et al., *Abnormal Expressions of DNA Glycosylase Genes NEIL1, NEIL2, and NEIL3 Are Associated with Somatic Mutation Loads in Human Cancer*. *Oxidative medicine and cellular longevity*, 2016. **2016**(undefined): p. 1546392.
11. Huang, W., *Regulation of the expression of the DNA repair endonuclease 8-like 3 gene in hepatocellular carcinoma*. 2019, AACR.
12. Ikeda, M., et al., *Phase 1 studies of 3-and 6-peptide vaccine cocktail derived from GPC3, WDRPUH, and NEIL3 in patients with advanced hepatocellular carcinoma (HCC)*. 2018, American Society of Clinical Oncology.

13. Q, L., et al., *HCCDB: A Database of Hepatocellular Carcinoma Expression Atlas*. Genomics, proteomics & bioinformatics, 2018. **16**(4): p. 269-275.
14. L, Z., et al., *Prognosis Analysis of Histone Deacetylases mRNA Expression in Ovarian Cancer Patients*. Journal of Cancer, 2018. **9**(23): p. 4547-4555.
15. GX, H., et al., *Mining expression and prognosis of topoisomerase isoforms in non-small-cell lung cancer by using OncoPrint and Kaplan-Meier plotter*. PloS one, 2017. **12**(3): p. e0174515.
16. Chandrashekar, D.S., et al., *UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses*. Neoplasia, 2017. **19**(8): p. 649-658.
17. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles*. Proceedings of the National Academy of Sciences, 2005. **102**(43): p. 15545-15550.
18. Vasaikar, S.V., et al., *LinkedOmics: analyzing multi-omics data within and across 32 cancer types*. Nucleic acids research, 2017. **46**(D1): p. D956-D963.
19. Li, T., et al., *TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells*. Cancer Research, 2017. **77**(21): p. e108.
20. Bo, L., et al., *Comprehensive analyses of tumor immunity: implications for cancer immunotherapy*. Genome Biology, 2016. **17**(1): p. 174.
21. Aran, D., M. Sirota, and A.J. Butte, *Systematic pan-cancer analysis of tumour purity*. Nature Communications, 2015. **6**: p. 8971.
22. Shahamatdar, S., et al., *Germline features associated with immune infiltration in solid tumors*. bioRxiv, 2019: p. 586081.
23. Givechian, K.B., et al., *Identification of an immune gene expression signature associated with favorable clinical features in Treg-enriched patient tumor samples*. NPJ genomic medicine, 2018. **3**(1): p. 14.
24. Danaher, P., et al., *Gene expression markers of tumor infiltrating leukocytes*. Journal for immunotherapy of cancer, 2017. **5**(1): p. 18.
25. Siemers, N.O., J.L. Holloway, and H. Chang, *Genome-wide association analysis identifies genetic correlates of immune infiltrates in solid tumors*. Plos One, 2017. **12**(7): p. e0179726.
26. Sofia, S. and M. Jorma, *The role of tumour-associated macrophages in bone metastasis*. Journal of Bone Oncology, 2016. **5**(3): p. 135-138.
27. Z, T., et al., *GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses*. Nucleic acids research, 2017. **45**(null): p. W98-W102.
28. Zhang, H., et al., *Analysis of loss of heterozygosity on chromosome 4q in hepatocellular carcinoma using high-throughput SNP array*. Oncology reports, 2010. **23**(2): p. 445-455.
29. Mattar, M.-A.M., et al., *Polymorphisms of base-excision repair genes and the hepatocarcinogenesis*. Gene, 2018. **675**: p. 62-68.
30. Chen, X., et al., *Interplay between base excision repair protein XRCC1 and ALDH2 predicts overall survival in lung and liver cancer patients*. Cellular Oncology, 2018. **41**(5): p. 527-539.

31. M, D., et al., *Dynamic modeling of signal transduction by mTOR complexes in cancer*. Journal of theoretical biology, 2019. **483**(undefined): p. 109992.
32. Shao, X., et al., *Reduced expression of Nrdp1 predicts a poor prognosis in human hepatocellular carcinoma*. OncoTargets and therapy, 2018. **11**: p. 4955.
33. Jia, M., et al., *lincRNA-p21 inhibits invasion and metastasis of hepatocellular carcinoma through Notch signaling-induced epithelial–mesenchymal transition*. Hepatology Research, 2016. **46**(11): p. 1137-1144.
34. Liu, Z., et al., *Reduced expression of cenp-e in human hepatocellular carcinoma*. Journal of experimental & clinical cancer research, 2009. **28**(1): p. 156.
35. Zhang, D., et al., *MiR-3613-3p affects cell proliferation and cell cycle in hepatocellular carcinoma*. Oncotarget, 2017. **8**(54): p. 93014.
36. Liu, W., et al., *Overexpression of non-SMC condensin I complex subunit G serves as a promising prognostic marker and therapeutic target for hepatocellular carcinoma*. International journal of molecular medicine, 2017. **40**(3): p. 731-738.
37. G, H., et al., *FOX M1 promotes hepatocellular carcinoma progression by regulating KIF4A expression*. Journal of experimental & clinical cancer research : CR, 2019. **38**(1): p. 188.
38. Liu, M., S. Doubl  , and S.S. Wallace, *Neil3, the final frontier for the DNA glycosylases that recognize oxidative damage*. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 2013. **743**: p. 4-11.
39. Liu, Z., Q. Sun, and X. Wang, *PLK1, a potential target for cancer therapy*. Translational oncology, 2017. **10**(1): p. 22-32.
40. Herrera, M.C., et al., *Cdk1 gates cell cycle-dependent tRNA synthesis by regulating RNA polymerase III activity*. Nucleic acids research, 2018. **46**(22): p. 11698-11711.
41. Fang, H., et al., *RecQL4-Aurora B kinase axis is essential for cellular proliferation, cell cycle progression, and mitotic integrity*. Oncogenesis, 2018. **7**(9): p. 68.
42. Chen, Y.-L., et al., *The E2F transcription factor 1 transactivates stathmin 1 in hepatocellular carcinoma*. Annals of surgical oncology, 2013. **20**(12): p. 4041-4054.
43. Chang, L., et al., *SIRT5 promotes cell proliferation and invasion in hepatocellular carcinoma by targeting E2F1*. Molecular medicine reports, 2018. **17**(1): p. 342-349.
44. Chen, H., et al., *CDCA5, Transcribed by E2F1, Promotes Oncogenesis by Enhancing Cell Proliferation and Inhibiting Apoptosis via the AKT Pathway in Hepatocellular Carcinoma*. Journal of Cancer, 2019. **10**(8): p. 1846.
45. Neurauter, C.G., L. Luna, and M. Bj  r  s, *Release from quiescence stimulates the expression of human NEIL3 under the control of the Ras dependent ERK–MAP kinase pathway*. DNA repair, 2012. **11**(4): p. 401-409.
46. T, Y., et al., *Programmed cell death 1 forms negative costimulatory microclusters that directly inhibit T cell receptor signaling by recruiting phosphatase SHP2*. The Journal of experimental medicine, 2012. **209**(6): p. 1201-17.

47. Ohue, Y. and H. Nishikawa, *Regulatory T (Treg) cells in cancer: Can Treg cells be a new therapeutic target?* *Cancer science*, 2019. **110**(7): p. 2080.
48. Huang, Y.-H., et al., *CEACAM1 regulates TIM-3-mediated tolerance and exhaustion*. *Nature*, 2015. **517**(7534): p. 386.
49. Torisu, K., et al., *Hematopoietic Tissue-Specific Expression of Mouse Neil3 for Endonuclease VIII-Like Protein*. *Journal of biochemistry*, 2005. **138**(6): p. 763-772.
50. Jørgensen, S.F., B. Fevang, and P. Aukrust, *Autoimmunity and inflammation in CVID: a possible crosstalk between immune activation, gut microbiota, and epigenetic modifications*. *Journal of clinical immunology*, 2019. **39**(1): p. 30-36.

Supplemental Figures

Supplementary Figure 1. Gene expression correlation analysis for NEIL3, CENPE, NCAPG, and KIF4A (LinkedOmics). The scatter plot shows Spearman correlation of NEIL3 expression with expression of (A) CNEPE, (B) NVAPG, and (C) KIF4A.

Supplementary Figure 2. Protein-protein interaction network transcription factor E2F1_Q6-target networks (GeneMANIA). Protein-protein interaction (PPI) network and functional analysis indicating the gene set that was enriched in the target network of transcription factor E2F1_Q6. Different colors of the network edge indicate the bioinformatics methods applied: co-expression, website prediction, pathway, physical interactions, shared protein domains, genetic interaction and co-localization. The different colors for the network nodes indicate the biological functions of the set of enrichment genes.

Figures

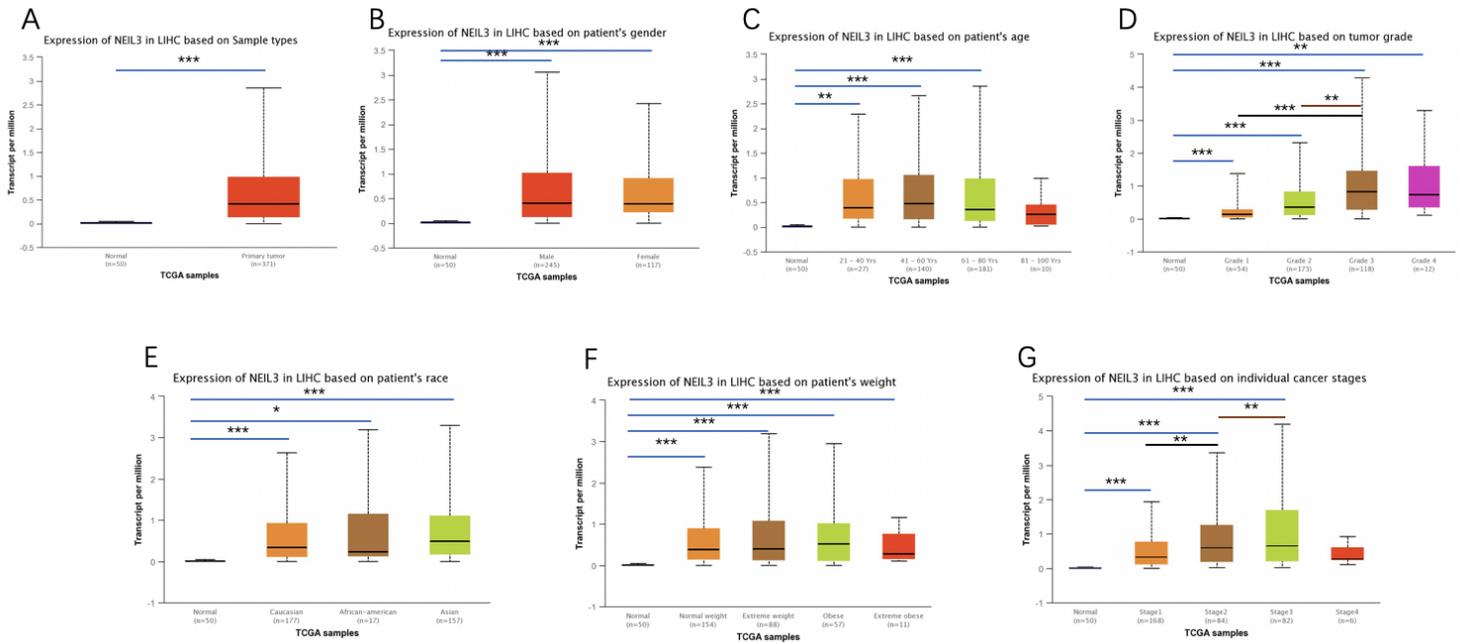


Figure 6

NEIL3 transcription in subgroups of patients with hepatocellular carcinoma, stratified based on gender, age and other criteria (UALCAN). (A) Boxplot showing relative expression of NEIL3 in normal and LHC samples. (B) Boxplot showing relative expression of NEIL3 in normal individuals of either gender or male or female LHC patients. (C) Boxplot showing relative expression of NEIL3 in normal individuals of any age or in LHC patients aged 21–40, 41–60, 61–80, or 81–100 yr. (D) Boxplot showing relative expression of RBM8A in normal individuals or LHC patients with grade 1, 2, 3 or 4 tumors. (E) Boxplot showing relative expression of NEIL3 in normal individuals of any ethnicity or in LHC patients of Caucasian, African-American or Asian ethnicity. (F) Boxplot showing relative expression of NEIL3 in normal individuals or in LHC patients with different weight. (G) Boxplot showing relative expression of NEIL3 in normal individuals or in LHC patients in stages 1, 2, 3 or 4. Data are mean \pm SE. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

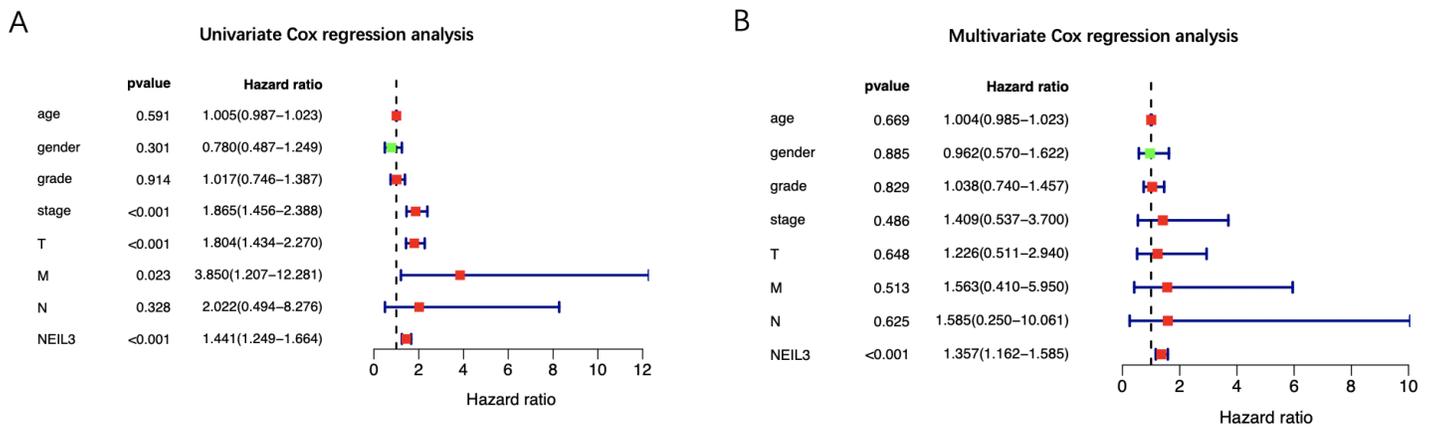


Figure 8

Univariate (A) and multiple (B) regression analysis of hepatocellular carcinoma and the relationships between the age, gender, grade, stage, T stage, distant metastasis, lymph node metastasis and expression of NEIL3. Green dot means hazard ratio (HR) median value is less than 1, red dot means HR median value is greater than 1.

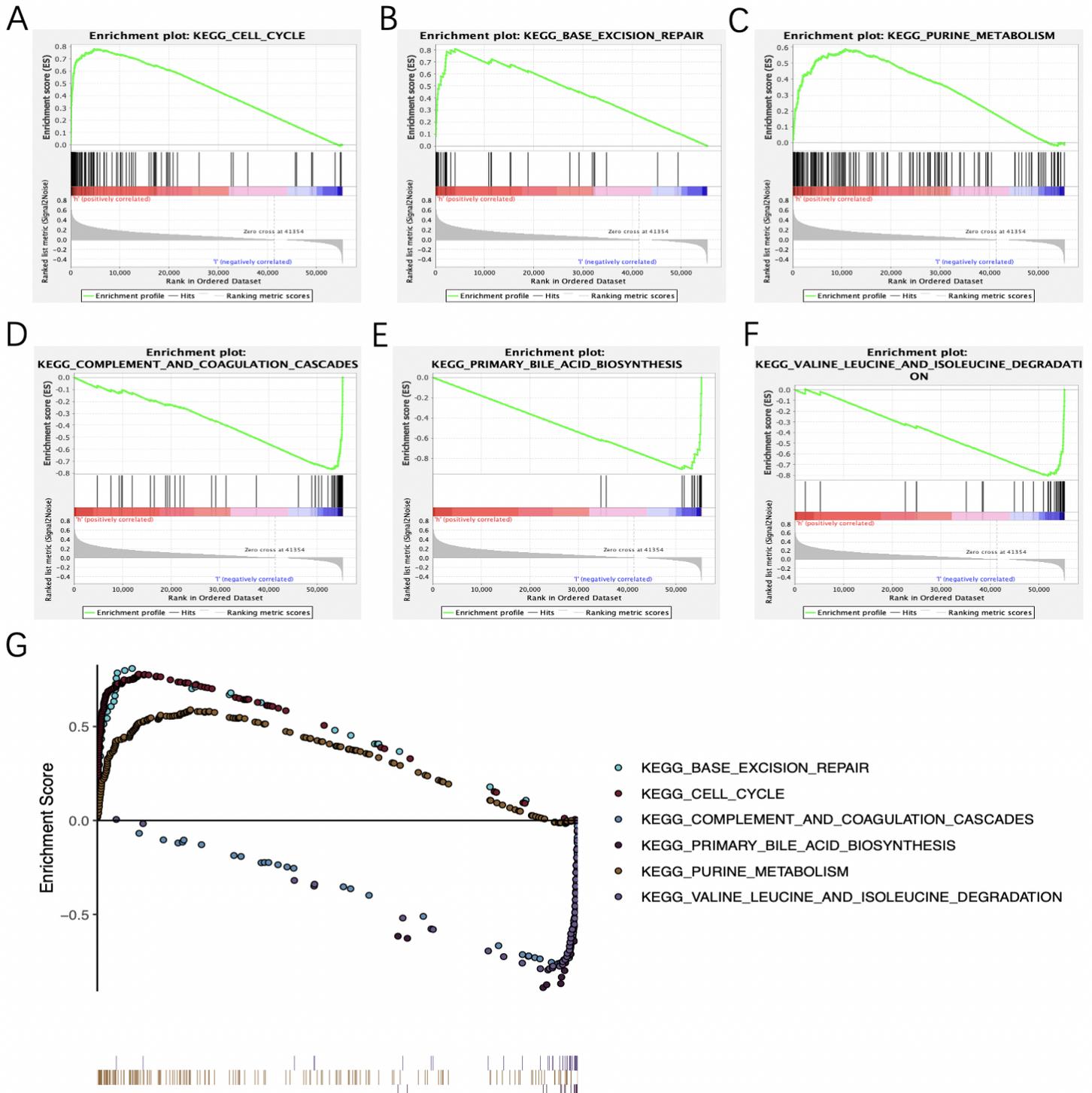


Figure 10

Enrichment plots from gene set enrichment analysis (GSEA). GSEA results showing (A) Cell cycle, (B) Base excision repair, (C) Purine metabolism are positively differentially enriched in high NEIL3 expression group;

while (D) Complement and coagulation cascades, (E) Primary bile acid biosynthesis, (F) valine leucine and isoleucine degradation are positively differentially enriched in low NEIL3 expression group.

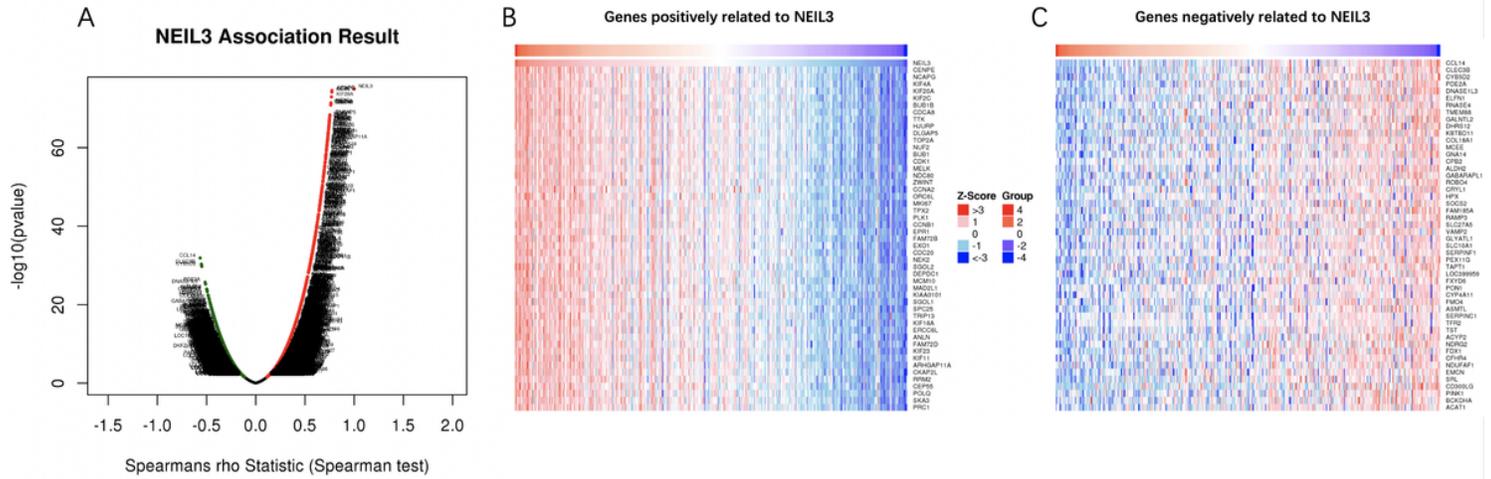


Figure 12

Genes differentially expressed in correlation with NEIL3 in hepatocellular carcinoma (LinkedOmics). (A) A Spearman test was used to analyze correlations between NEIL3 and genes differentially expressed in LIHC. (B–C) Heat maps showing genes positively and negatively correlated with NEIL3 in LIHC (TOP 50). Red indicates positively correlated genes and green indicates negatively correlated genes.

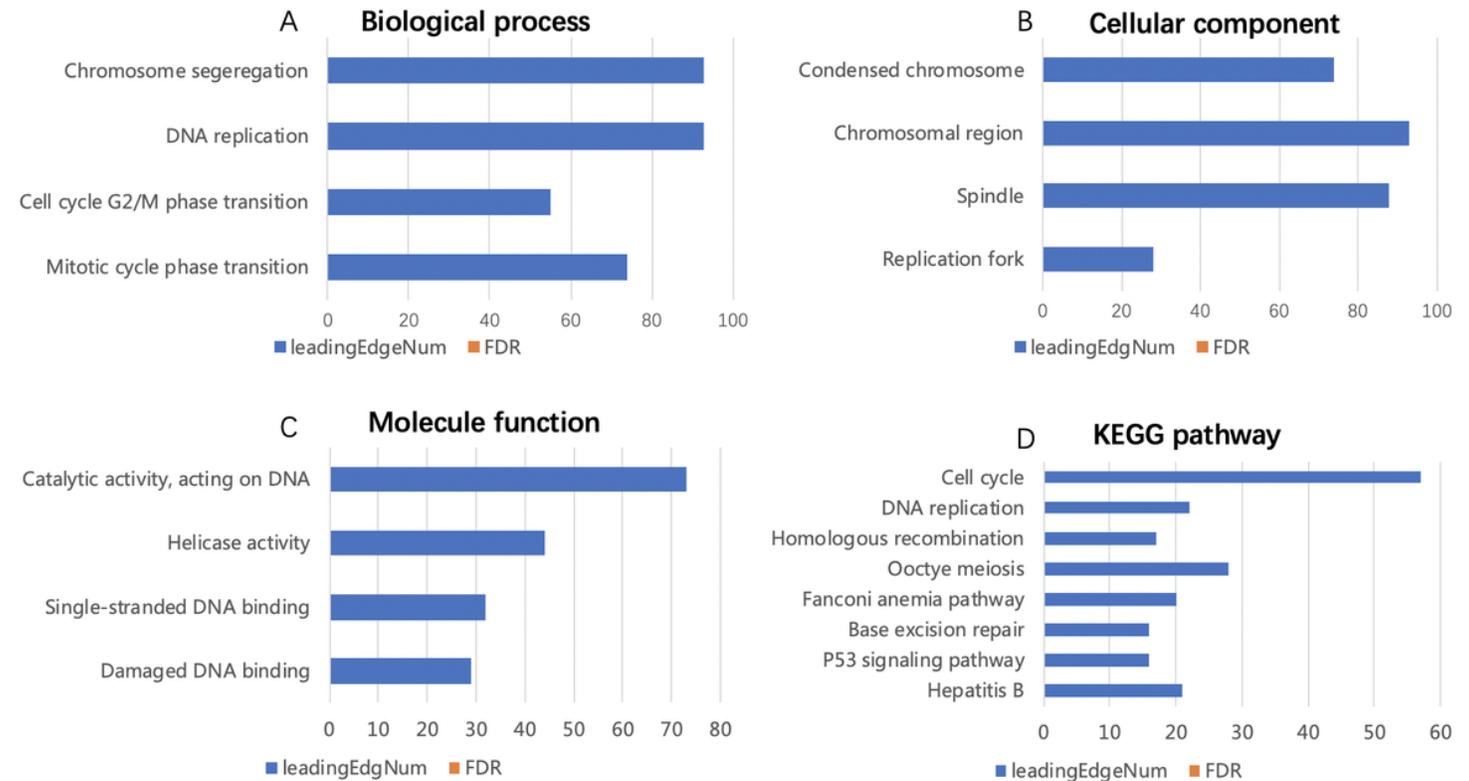
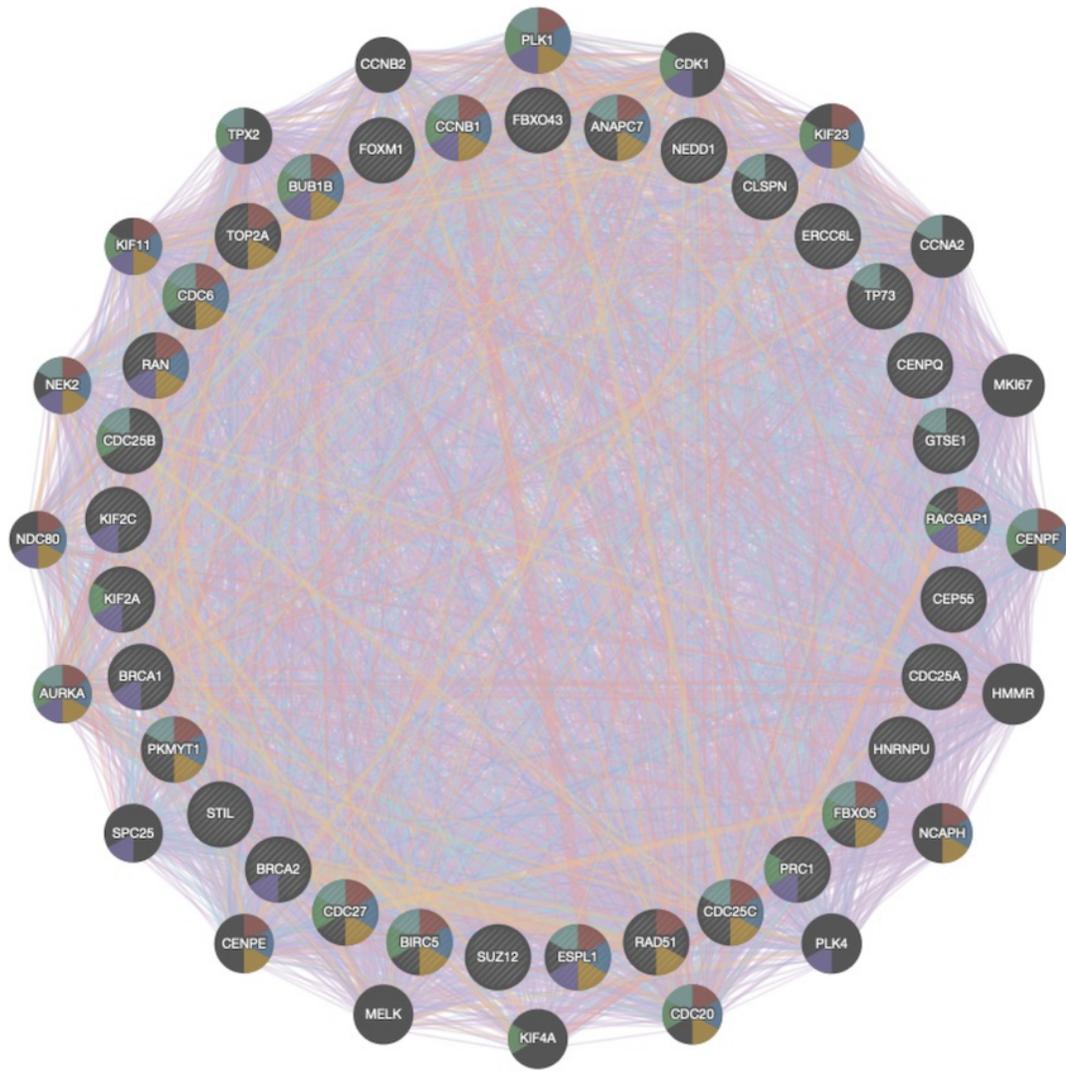


Figure 14

Significantly enriched GO annotations and KEGG pathways of NEIL3 in hepatocellular carcinoma. The significantly enriched GO annotations and KEGG pathways of NEIL3 co-expression genes in LIHC were

analyzed using GSEA. (A) Biological processes. (B) Cellular components. (C) Molecular functions. (D) KEGG pathway analysis. The blue column represents the LeadingEdgeNum, and the orange represents the false discovery rate (FDR). The FDR from GSEA in the figure is 0.



Networks

- Co-expression
- Co-localization
- Physical Interactions
- Predicted
- Pathway
- Shared protein domains
- Genetic Interactions

Functions

- nuclear division
- mitosis
- organelle fission
- microtubule cytoskeleton organization
- spindle
- regulation of mitotic cell cycle

Figure 16

Protein-protein interaction network of PLT1 kinase-target networks (GeneMANIA). Protein-protein interaction (PPI) network and functional analysis indicating the gene set that was enriched in the target network of PLT1 kinases. Different colors of the network edge indicate the bioinformatics methods applied: co-

expression, website prediction, pathway, physical interactions, shared protein domains, genetic interaction and co-localization. The different colors for the network nodes indicate the biological functions of the set of enrichment genes.

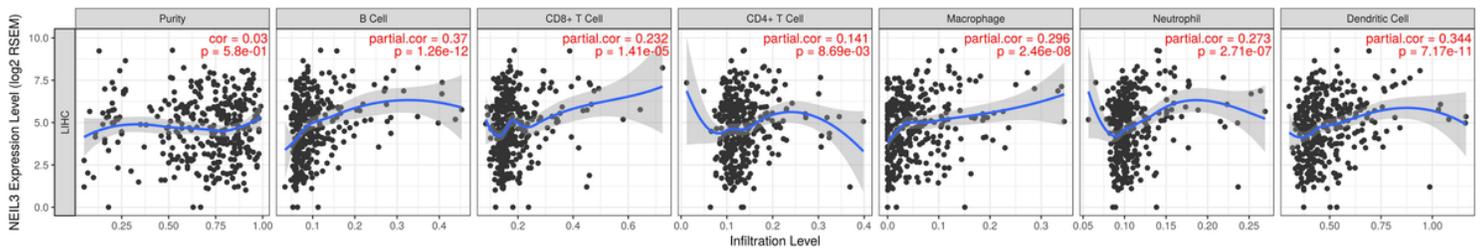


Figure 17

Correlation of NEIL3 expression with immune infiltration level in hepatocellular carcinoma. Expression has significant positive correlations with infiltrating levels of B cells, CD8+ T cells, CD4+ T cells, macrophages, neutrophils, and dendritic cells, other than tumor purity.

Supplementary Files

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

- [SupplementaryTable3.docx](#)
- [SupplementaryTable3.docx](#)
- [SuppFig2.tiff](#)
- [supplementaryfig1.tiff](#)
- [SuppFig2.tiff](#)
- [SupplementaryTable1.docx](#)
- [SupplementaryTable2.docx](#)
- [SupplementaryTable1.docx](#)
- [SupplementaryTable2.docx](#)
- [supplementaryfig1.tiff](#)