

# Potential value of PRKDC as a therapeutic target and prognostic biomarker in pan-cancer

Xiawei Yang<sup>a</sup>, Feng Yang<sup>b</sup>, Liugen Lan<sup>c,d,e,f</sup>, Ning Wen<sup>c,d,e,f</sup>, Haibin Li<sup>c,d,e,f</sup>, Xuyong Sun<sup>c,d,e,f,\*</sup> 

## Abstract

**Background:** While protein kinase, DNA-activated, catalytic subunit (PRKDC) plays an important role in double-strand break repair to retain genomic stability, there is still no pan-cancer analysis based on large clinical information on the relationship between PRKDC and different tumors. For the first time, this research used numerous databases to perform a pan-cancer review for PRKDC to explore the possible mechanism of PRKDC in the etiology and outcomes in various tumors.

**Methods:** PRKDC's expression profile and prognostic significance in pan-cancer were investigated based on various databases and online platforms, including TIMER2, GEPIA2, cBioPortal, CPTAC, and SangerBox. We applied the TIMER to identify the interlink of PRKDC and the immune infiltration in assorted tumors, and the SangerBox online platform was adopted to find out the relevance between PRKDC and immune checkpoint genes, tumor mutation burden, and microsatellite instability in tumors. GeneMANIA tool was employed to create a protein–protein interaction analysis, gene set enrichment analysis was conducted to perform gene enrichment analysis.

**Results:** Overall, tumor tissue presented a higher degree of PRKDC expression than adjacent normal tissue. Meanwhile, patients with high PRKDC expression have a worse prognosis. PRKDC mutations were present in almost all The Cancer Genome Atlas tumors and might lead to a better survival prognosis. The PRKDC expression level was shown a positive correlation with tumor-infiltrating immune cells. PRKDC high expression cohorts were enriched in “cell cycle” “oocyte meiosis” and “RNA-degradation” signaling pathways.

**Conclusions:** This study revealed the potential value of PRKDC in tumor immunology and as a therapeutic target and prognostic biomarker in pan-cancer.

**Abbreviations:** ACC = adrenocortical carcinoma, BLCA = bladder urothelial carcinoma, BRCA = breast invasive carcinoma, CESC = cervical squamous cell carcinoma and endocervical adenocarcinoma, CHOL = cholangiocarcinoma, CNA = copy number alteration, COAD = colon adenocarcinoma, CPTAC = clinical proteomic tumor analysis consortium, DC = dendritic cells, DFI = disease-free interval, DFS = disease-free survival, DLBC = lymphoid neoplasm diffuse large B-cell lymphoma, DNA-PKcs = DNA-dependent protein kinase catalytic subunit, DNA-PK = DNA-dependent protein kinase complex, DSB = double-strand break, DSS = disease-specific survival, ESCA = esophageal carcinoma, FC = fold change, GBM = glioblastoma multiforme, GEPIA 2 = gene expression profiling interactive analysis, version 2, GSEA = gene set enrichment analysis, GTEx = genotype-tissue expression, HNSC = head and neck squamous cell carcinoma, HR = homologous recombination, ICI = immune checkpoint inhibitor, KEGG = Kyoto encyclopedia of genes and genomes, KICH = kidney chromophobe, KIRC = kidney renal clear cell carcinoma, KIRP = kidney renal papillary cell carcinoma, LAML = acute myeloid leukemia, LGG = brain lower grade glioma, LIHC = liver hepatocellular carcinoma, LUAD = lung adenocarcinoma, LUSC = lung squamous cell carcinoma, MESO = mesothelioma, MMR = mismatch repair, MSI = microsatellite instability, NHFJ = nonhomologous final junction, OS = overall survival, OV = ovarian serous cystadenocarcinoma, PAAD = pancreatic adenocarcinoma, PCPG = pheochromocytoma and paraganglioma, PFI = progression-free interval, PFS = progression-free survival, PI3K = phosphatidylinositol 3-kinase-like, PPI = protein–protein interaction, PRAD = prostate adenocarcinoma, PRKDC = protein kinase, DNA-activated, catalytic subunit, RCC = renal cell carcinoma, READ = rectum adenocarcinoma, SARC = sarcoma, SKCM = skin cutaneous melanoma, STAD = stomach adenocarcinoma, TCGA = The Cancer Genome Atlas, TGCT = testicular germ cell tumors, THCA = thyroid carcinoma, THYM = thymoma, TIICs = tumor-infiltrating immune cells, TIMER2.0 = Tumor Immune Estimation Resource, version 2, TMB = tumor mutation burden, TPM = Transcripts per million, UCEC = uterine corpus endometrial carcinoma, UCS = uterine carcinosarcoma, UVM = uveal melanoma.

**Keywords:** DNA-PKcs, genomic alteration, immune infiltration, pan-cancer, PRKDC, prognosis, protein phosphorylation, tumor microenvironment

This work was supported by the National Natural Science Foundation of China (81670596).

The authors have no conflicts of interest to disclose.

This manuscript was previously posted to Research Square: doi: 10.21203/rs.3.rs-789778/v1.

The raw data of this study are derived from the GTEx (<https://gtexportal.org/home/>) database and the TCGA (<https://www.cancer.gov/>) database, which are publicly available databases.

As the data for this study were obtained from public databases, ethical approval was not required.

Supplemental Digital Content is available for this article.

<sup>a</sup> Guangxi Medical University, Nanning, Guangxi Zhuang Autonomous Region, China, <sup>b</sup> Department of Gynecology, The Second Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi Zhuang Autonomous Region, China, <sup>c</sup> Transplant Medical Center, The Second Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi Zhuang Autonomous Region, China, <sup>d</sup> Guangxi Key Laboratory of Organ Donation and Transplantation, Nanning, Guangxi Zhuang Autonomous Region, China, <sup>e</sup> Guangxi Key Laboratory for Transplantation Medicine, Nanning, Guangxi Zhuang Autonomous Region, China, <sup>f</sup> Guangxi Transplantation Medicine Research Center of Engineering Technology, Nanning, Guangxi Zhuang Autonomous Region, China.

## 1. Introduction

Cells encounter many DNA lesions every day that imperil their genomic completeness, and the most toxic one is double-strand break (DSB). As we all know, one unrepaired DSB might induce cell death, rather a misrepaired DSB can lead to chromosomal abnormalities (e.g., deletions, translocations, and fusions), which may contribute to a loss of heterozygosity, induce genetic instability, neoplastic transformation, and ultimately, cancer.

Organisms have developed wrought genomic stability maintenance systems to prevent cancer<sup>[1]</sup> by identifying broken DNA sites and restoring the DNA damage.<sup>[2]</sup> In human cells, DSB repair is mediated by 2 distinguished ways: nonhomologous final junction (NHFJ) and homologous recombination (HR), and NHEJ was considered quicker and more efficient than HR.<sup>[3]</sup>

DNA-dependent protein kinase catalytic subunit (DNA-PKcs) or serine/threonine protein kinase catalytic subunit are the other names for PRKDC. It belongs to the phosphatidylinositol 3-kinase-like (PIKK) family and presents at approximately all mammalian cells.<sup>[4]</sup> PRKDC interacts with the Ku70/Ku80 heterodimer, then together join into the ligation step of the NHEJ process and, ultimately, helps to preserve genomic integrity.<sup>[5]</sup>

PRKDC has recently attracted a lot of interest in being a therapeutic target and promising biomarker for many human cancers,<sup>[6–13]</sup> but a pan-cancer investigation of the correlation between PRKDC with various types of tumors based on these large-scale clinical data is currently lacking. In this study, for the first time, we conducted a pan-cancer analysis of PRKDC based on a variety of databases and online platforms to investigate the latent molecular processes of PRKDC in different tumor's pathogenesis and clinical outcomes.

## 2. Materials and Methods

### 2.1. PRKDC expression analysis

By using Tumor Immune Estimation Resource, version 2 (TIMER2.0, <http://timer.cistrome.org/>) database,<sup>[14]</sup> the expression levels of PRKDC in tumor tissues compared with neighboring normal tissue were analyzed in all The Cancer Genome Atlas (TCGA, <https://www.cancer.gov/>) tumors. For those tumors without normal tissues, we used the Gene Expression Profiling Interactive Analysis, version 2 (GEPIA 2, <http://gepia2.cancer-pku.cn/>)<sup>[15]</sup> to generate boxplots from the Genotype-Tissue Expression (GTEx, <https://gtexportal.org/home/>) database, and selected “Match TCGA normal and GTEx data” pattern. This study furthermore generated violin plots of PRKDC expression according to pathological stages of all TCGA tumors by using the “Pathological Stage Plot” modular in GEPIA 2, with cutoff values set to  $P$  value = .01 and  $\log_2$  FC (fold change) = 1. Transcripts per million (TPM) values were transformed into a  $\log_2$  (TPM + 1).

The UALCAN (<http://ualcan.path.uab.edu/>)<sup>[16]</sup> enabled us to carried out phosphorylation investigations of PRKDC protein across various tumors from the Clinical proteomic tumor analysis consortium (CPTAC) database.<sup>[17]</sup> This work compared PRKDC total protein and phosphoprotein expression levels in 6 tumors and neighbor normal tissues, including breast cancer, ovarian cancer, colon cancer, clear cell renal cell carcinoma (RCC), uterine corpus endometrial carcinoma (UCEC), and lung adenocarcinoma (LUAD).

### 2.2. Survival analysis

Connection between PRKDC expression and survival of all TCGA cancers was investigated using “Survival Map” and “Survival Analysis” modules of GEPIA 2. To separate the high and low PRKDC expression groups, a 50% cutoff value was chosen. SangerBox online website (<http://sangerbox.com>) is a comprehensive Chinese bioinformatics analysis platform. The “Gene-KM plotter” module of SangerBox was used to compare the overall survival (OS), disease-free interval (DFI), disease-specific survival (DSS), and progression-free interval (PFI) in high and low expression cohorts. Hazard ratios (HRs) were computed using a 95% confidence interval and a  $P$  value.

### 2.3. Genetic alteration analysis

We acquired PRKDC mutation status in tumor patients from the cBioPortal (<http://www.cbioportal.org/>) platform.<sup>[18]</sup> The “Cancer Types Summary” section displayed the frequency of alteration, types of mutation, and copy number alteration data. The “Mutations” module presented the mutation site information in the PRKDC protein structural schematic diagram. Using the “Comparison/Survival” module, Kaplan–Meier plots were created to compare DSS, OS, DFS, and PFS in cohorts with or without PRKDC genetic mutation, with log-rank  $P$  values.

### 2.4. Immunological analysis

To observe the interaction between tumor cells and immune cells, the Tumor Immune Estimation Resource (TIMER, <https://cistrome.shinyapps.io/timer/>) database<sup>[19]</sup> was employed. By using “Gene” module of TIMER, we determined the tumors' purity and discovered the relationship between PRKDC and the quantity of 6 tumor-infiltrating immune cells (TIICs) subsets in 39 kinds of tumors with Spearman correlations. CD4+ T cells, CD8+ T cells, B cells, neutrophils, dendritic cells (DCs), and macrophages are among these 6 categories of TIICs. The correlations between PRKDC and 28 subtypes of immune cell were determined and evaluated via the “Gene-Immune Analysis” module of SangerBox. We also looked at the relational between PRKDC and 47 kinds of gene markers of TIICs in distinct tumors.<sup>[20]</sup>

The all somatic nonsynonymous mutation counts per megabase in coding sequence were used to calculate the tumor mutation burden (TMB), which may be a promising marker for predicting immunotherapy response.<sup>[21]</sup> Microsatellite instability (MSI) is a molecular tumor characteristic characterized by spontaneous nucleotide loss or gain on short tandem repetitive DNA sequences, and is the result of a defective DNA mismatch repair (MMR) system.<sup>[22]</sup> Analysis and visualization of TMB and MSI in different tumors were carried out using SangerBox online platform.

### 2.5. Protein–protein interaction network and enrichment analysis

We built a protein–protein interaction (PPI) network for PRKDC using the GeneMANIA online platform (<https://genemania.org/>)<sup>[23]</sup> to find out the mechanism of PRKDC in tumorigenesis. The first 5 terms of KEGG (Kyoto encyclopedia of genes and genomes) and HALLMARK analyses were displayed in the high/low PRKDC expression cohorts, respectively, to discover

\* Correspondence: Xuyong Sun, No. 166, Daxuedong Road, Nanning, Guangxi Zhuang Autonomous Region, 530007, China (e-mail: [sxywnzyx@163.com](mailto:sxywnzyx@163.com)).

Copyright © 2022 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the Creative Commons Attribution License 4.0 (CCBY), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

How to cite this article: Yang X, Yang F, Lan L, Wen N, Li H, Sun X. Potential value of PRKDC as a therapeutic target and prognostic biomarker in pan-cancer. *Medicine* 2022;101:27(e29628).

Received: 10 November 2021 / Received in final form: 19 February 2022 / Accepted: 6 May 2022

<http://dx.doi.org/10.1097/MD.00000000000029628>

the biological signaling pathway.  $|\text{NES}| > 1$ ,  $P < .01$ , and  $\text{FDR} < 0.25$  were defined as significant gene sets.<sup>[24]</sup>

### 3. Results

#### 3.1. PRKDC expression analysis results

We adopted the TCGA and GTEx datasets to look at PRKDC expression levels across diverse tumors. The expression of PRKDC in the tumor tissues of cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC), brain lower grade glioma (LGG), liver hepatocellular carcinoma (LIHC), LUAD, lung squamous cell carcinoma (LUSC), bladder urothelial carcinoma (BLCA), breast invasive carcinoma (BRCA), cholangiocarcinoma, lymphoid neoplasm diffuse large B-cell lymphoma, rectumadenocarcinoma, stomach adenocarcinoma (STAD), thymoma (THYM), uterine carcinosarcoma, colon adenocarcinoma (COAD), esophageal carcinoma (ESCA), and head and neck squamous cell carcinoma (HNSC) were higher compare with those in neighboring normal tissues, as illustrated in Fig. 1A and B. By contrast, PRKDC expression levels were significantly lower in tumor tissues of kidney renal clear cell carcinoma (KIRC), kidney renal papillary cell carcinoma (KIRP), and thyroid carcinoma (THCA). But, there was no significant difference in other tumors, including glioblastoma multiforme (GBM), kidney chromophobe (KICH), pancreatic adenocarcinoma (PAAD), pheochromocytoma and paraganglioma, prostate adenocarcinoma, UCEC, adrenocortical carcinoma (ACC), acute myeloid leukemia, ovarian serous cystadenocarcinoma (OV), sarcoma (SARC), skin cutaneous melanoma (SKCM), and testicular germ cell tumors (all  $P > .05$ ), as indicated in Figure 1A and Figure S1A, Supplemental Digital Content, <http://links.lww.com/MD/G879>.

We applied the “Pathological Stage Plot” modular of GEPIA 2 to confirmed a positive association between PRKDC expression and advanced cancer stage in OV, SKCM, and UCEC (all  $P < .05$ , Fig. 1C) but not the others (Fig. S1B, Supplemental Digital Content, <http://links.lww.com/MD/G879>). The results from the CPTAC dataset confirmed that PRKDC total protein is highly expressed in breast cancer, ovarian cancer, colon cancer, UCEC, and LUAD compared with normal controls, but not clear cell RCC (Fig. 1D).

#### 3.2. Survival analysis results

We explored the prognosis value of PRKDC in pan-cancer base on several databases and platform. First, we evaluated the association between PRKDC and prognosis (OS and DFS) using GEPIA 2. As shown in Figures 2A, it suggested a poorer OS in high PRKDC expression cohort rather than low PRKDC expression cohort of LIHC (HR = 1.6,  $P = .01$ ), ACC (HR = 2.8,  $P = .015$ ), LGG (HR = 1.9,  $P = .00099$ ), LUAD (HR = 1.5,  $P = .0045$ ), mesothelioma (MESO) (HR = 2.2,  $P = .002$ ), SARC (HR = 1.8,  $P = .0042$ ), and UVM (HR = 3.2,  $P = .023$ ). High PRKDC expression cohorts also exhibited shorter DFS than low PRKDC expression cohorts for ACC (HR = 2.6,  $P = .0069$ ), KIRP (HR = 2,  $P = .017$ ), LGG (HR = 1.5,  $P = .018$ ), MESO (HR = 2,  $P = .016$ ), and SARC (HR = 1.4,  $P = .048$ ).

Next, survival data from the SangerBox platform revealed a link between high PRKDC expression and bad OS (Fig. 2B) for LUAD ( $P = .015$ ), UCEC ( $P = .0037$ ), BLCA ( $P = .015$ ), PAAD ( $P = .019$ ), KIRP ( $P = .0067$ ), LIHC ( $P = .00038$ ), SARC ( $P = .0084$ ), MESO ( $P = .0031$ ), LGG ( $P = .000047$ ), KICH ( $P = .0023$ ), ACC ( $P = .007$ ) and UVM ( $P = .0015$ ), but a greater OS in KIRC ( $P = .049$ ) and rectumadenocarcinoma ( $P = .013$ ). High PRKDC expression cohort was related to a shorter DFI (Fig. 2C) in PAAD ( $P = .003$ ) and KIRP ( $P = .013$ ) when compared with a low PRKDC expression cohort. Similar to OS, high PRKDC expression group showed worse

DSS (Fig. 2D) compare with the low PRKDC expression group for LUAD ( $P = .013$ ), UCEC ( $P = .0071$ ), BLCA ( $P = .0048$ ), PAAD ( $P = .017$ ), KIRP ( $P = .0024$ ), LIHC ( $P = .0085$ ), SARC ( $P = .0099$ ), MESO ( $P = .002$ ), LGG ( $P = .000081$ ), KICH ( $P = .00077$ ), ACC ( $P = .0088$ ), and UVM ( $P = .00096$ ). Additionally, high PRKDC expression was relevant with shorter PFI for LUAD ( $P = .007$ ), UCEC ( $P = .0017$ ), PAAD ( $P = .012$ ), KIRP ( $P = .0094$ ), LIHC ( $P = .027$ ), MESO ( $P = .0078$ ), LGG ( $P = .0027$ ), KICH ( $P = .012$ ), ACC ( $P = .000076$ ), and UVM ( $P = .000083$ ).

Finally, we evaluated prognostic value of PRKDC in pan-cancer (OS and RFS) on Kaplan–Meier plotter database (<https://kmplot.com/analysis/>). The PRKDC level negatively correlated with the OS (Fig. S2, Supplemental Digital Content, <http://links.lww.com/MD/G879>) in following cancers: BLCA ( $P = .03$ ), BRCA ( $P = .025$ ), CESC ( $P = .003$ ), KIRP ( $P = .0095$ ), LIHC ( $P = .0067$ ), LUAD ( $P = .001$ ), PAAD ( $P = .0069$ ), SARC ( $P = .00015$ ), THCA ( $P = .0069$ ), and UCEC ( $P = .0001$ ), but was positively correlated with the OS in ESCA ( $P = .0058$ ), KIRC ( $P = .0012$ ), and THYM ( $P = .021$ ). There was a negative relation between the PRKDC level and relapse-free survival (Fig. S3, Supplemental Digital Content, <http://links.lww.com/MD/G879>) in KIRP ( $P = .013$ ), LIHC ( $P = .029$ ), PAAD ( $P = .00038$ ), THCA ( $P = .015$ ), UCEC ( $P = .0061$ ), and SARC ( $P = .011$ ), but a positive correlation was shown in ESCA ( $P = .0075$ ) and OV ( $P = .011$ ).

The above data revealed that PRKDC is elevated in most human tumors, suggesting that it may serve as a biomarker for poor prognosis, though varied across different types of tumors.

#### 3.3. PRKDC alteration analysis results

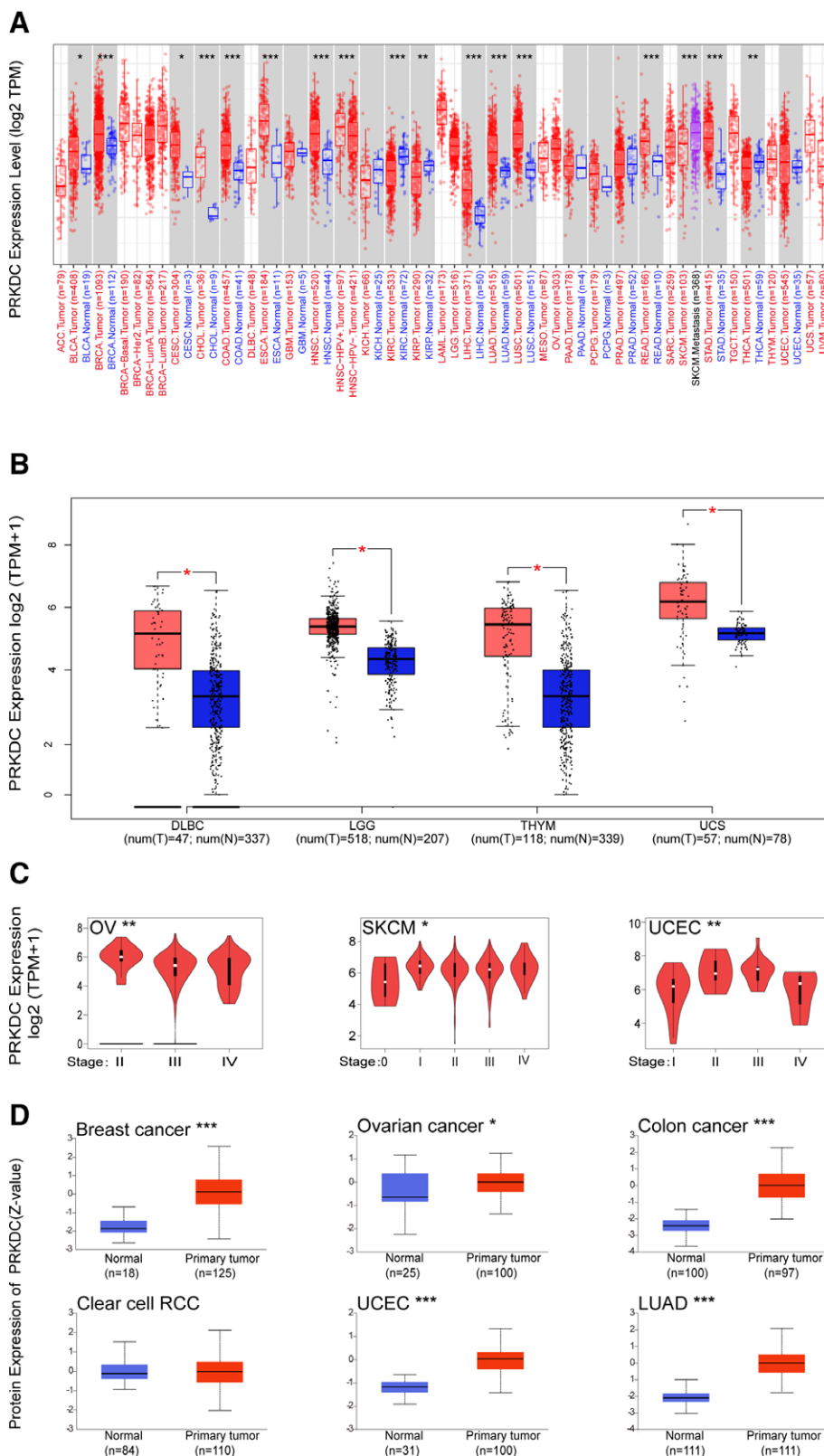
Because genomic mutations are linked to tumor oncogenesis and progression, we observed 10,953 patients from 32 TCGA studies in total (<https://www.cbiportal.org/>), and found that 804 (7%) of them carried at least one PRKDC gene mutation. As it is indicated in Figure 3A, “mutation” showed the highest frequency of PRKDC alteration in most tumors, including UCEC, SKCM, STAD, LIHC, LUAD, COAD, BLCA, HNSC, ESCA, LUSC, CESC, SARC, MESO, KIRP, KIRC, ACC, GBM, LGG, THYM, and THCA. In particular, “mutation” was the only genetic alteration type in KIRP (~2.5% frequency), ACC (~2% frequency), THYM (~1% frequency), and THCA (~1% frequency). PRKDC mutation frequency was the highest in UCEC (>18%), and the “amplification” alteration type was the primary type in uterine carcinosarcoma, with an alteration frequency nearby 15%.

The alteration types, mutated sites, and mutation cases numbers were exhibited in a schematic representation of PRKDC protein structure in Figure 3B. It showed that “Truncating” is the most common sort of genetic alteration, and N3604Kfs\*3/Tfs\*48 alteration was detected in 6 cases of UCEC and 6 cases of STAD.

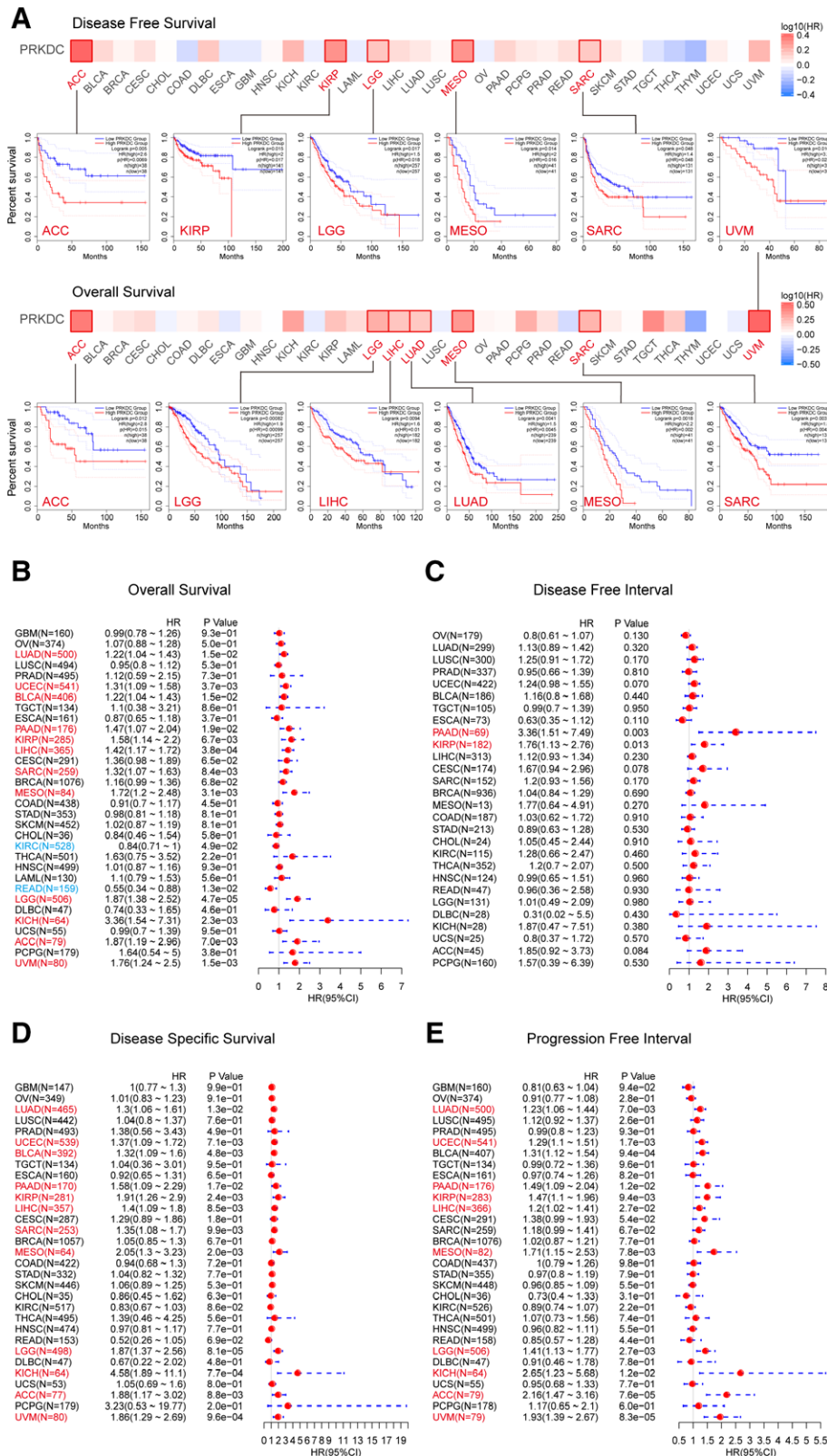
Additionally, we compared the survival prognosis in cohorts with or without PRKDC genetic alteration across different tumors. It indicated a better prognosis for PFS ( $P = 1.862e-03$ ), DSS ( $P = 4.566e-03$ ), and OS ( $P = 5.864e-03$ ) in UCEC cases (Fig. 3C) with altered PRKDC, and better prognosis for PFS ( $P = 1.582e-03$ ) and DSS ( $P = .0366$ ) in STAD cases (Fig. 3D) with altered PRKDC, compared with cases without PRKDC alteration.

#### 3.4. Protein phosphorylation analysis results

Using the CPTAC dataset, we evaluated the PRKDC phosphorylation level in tumor tissues as well normal tissues in 6 kinds of cancers, including breast cancer, colon cancer, clear cell RCC, UCEC, LUAD, and ovarian cancer. A schematic diagram summarizes the PRKDC phosphorylation sites and related tumors

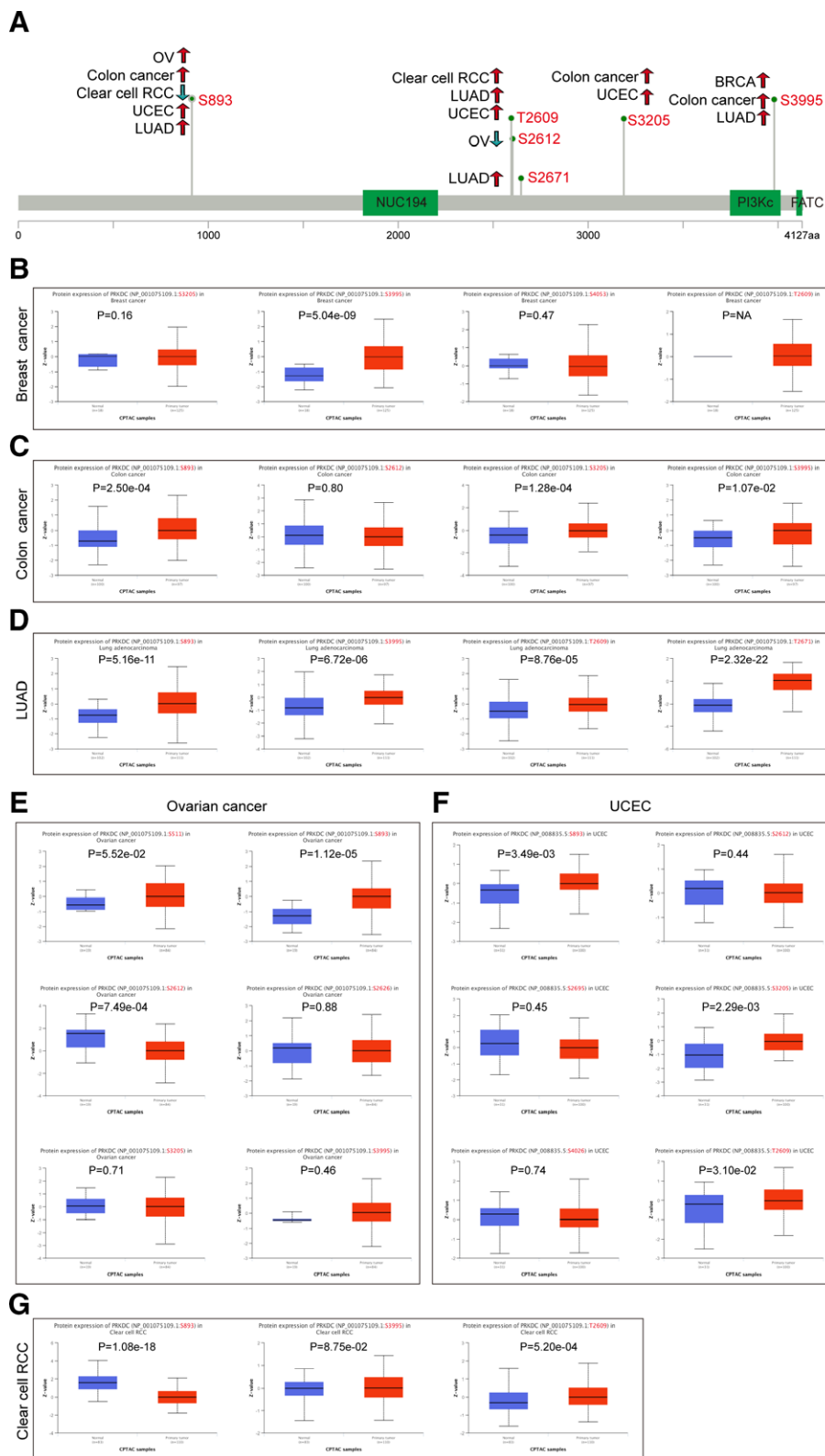


**Figure 1.** PRKDC expression in different cancers and pathological stages. (A) Comparison of PRKDC expression level in 33 tumor types and normal tissues determined by TIMER2. (B) For those tumors without normal tissues in TCGA database, we used GEPIA 2 to obtain box plots from the GTEx database, and positive results (DLBC, LGG, THYM, and UCS) were shown. T: tumor, N: normal. (C) Violin plots from GEPIA 2 confirmed a positive association between PRKDC expression and advanced cancer stages in OV, SKCM, and UCEC. TPM values were transformed into a log2(TPM+1) scale. (D) The results from the CPTAC dataset confirmed that PRKDC total protein is highly expressed in breast cancer, ovarian cancer, colon cancer, UCEC, and LUAD compared with normal controls, but not clear cell RCC. DLBC = lymphoid neoplasm diffuse large B-cell lymphoma, LGG = brain lower grade glioma, LUAD = lung adenocarcinoma, OV = ovarian serous cystadenocarcinoma, PRKDC = protein kinase, DNA-activated, catalytic subunit, RCC = renal cell carcinoma, SKCM = skin cutaneous melanoma, TIMER2.0 = Tumor Immune Estimation Resource, version 2, THYM = thymoma, UCEC = uterine corpus endometrial carcinoma, UCS = uterine carcinosarcoma. \* $P < .05$ ; \*\* $P < .01$ ; \*\*\* $P < .001$ .



**Figure 2.** Prognosis value of PRKDC in pan-cancer base on different platforms. (A) GEPIA 2 platform was adopted to carry out overall survival and disease-free survival analysis in all TCGA tumors. A 50% cutoff value was chosen to separating the high or low PRKDC expression groups. Survival maps and Kaplan-Meier plots with significant results were presented at the same time. Forest plots of overall survival (B), disease-free interval (C), disease-special survival (D), and progression-free interval (E) comparing high and low PRKDC expression cohorts were performed by SangerBox online website.  $P < .05$  is defined as significant. GEPIA 2 = gene expression profiling interactive analysis, version 2, PRKDC = protein kinase, DNA-activated, catalytic subunit, TCGA = The Cancer Genome Atlas.





**Figure 4.** Phosphorylation analysis of PRKDC protein across tumors. (A) The schematic diagram summarizes the PRKDC phosphorylation sites and related tumors. We compared the PRKDC phosphoprotein expression level between normal tissue and tumor tissue of breast cancer (B), colon cancer (C), LUAD (D), ovarian cancer (E), UCEC (F), and clear cell RCC (G) via the UALCAN.  $P < .05$  is defined as significant. LUAD = lung adenocarcinoma, PRKDC = protein kinase, DNA-activated, catalytic subunit, RCC = renal cell carcinoma, UCEC = uterine corpus endometrial carcinoma.

(Fig. 4A). The S893 locus of PRKDC exhibited an increased phosphorylation levels in colon cancer (Fig. 4C,  $P = 2.5e-04$ ), LUAD (Fig. 4D,  $P = 5.16e-11$ ), ovarian cancer (Fig. 4E,  $P = 1.12e-05$ ), and UCEC (Fig. 4F,  $P = 3.49 e-03$ ), but a decreased phosphorylation level in clear cell RCC (Fig. 4G,  $P = 1.08e-18$ ). Additionally, we confirmed that PRKDC phosphorylation of S893 was experimentally proved by one research<sup>[25]</sup> through the phosphoNET database (<http://www.phosphonet.ca/>). To learn more about the significance of S893 phosphorylation in carcinogenesis, further molecular investigations are needed.

### 3.5. Immune infiltration analysis results

TICs from tumor microenvironment (TME) are correlated with initiation, progression or metastasis of tumor,<sup>[26]</sup> so we explored the coefficient of TICs and PRKDC expression in diverse tumors of TCGA by using TIMER database. According to the findings, PRKDC was relevant to tumor purity in 9 different kinds of cancers, and also remarkably associated with the infiltration levels of B cells, CD8+ T cells, CD4+ T cells, macrophages, neutrophils, and DCs in 15, 21, 14, 18, 21, and 19 kinds of tumors, separately (Fig. 5A and Fig S4–7, Supplemental Digital Content, <http://links.lww.com/MD/G879>). In addition, KIRC, LIHC, and THCA show the most strongly association between PRKDC and immune infiltrating levels (Fig. 5B). For KIRC, a negative interrelation was confirmed between PRKDC gene and tumor purity ( $R = -0.105$ ,  $P = 2.41e-2$ ), but PRKDC expression and B cells ( $R = 0.295$ ,  $P = 1.15e-10$ ), CD8+ T cells ( $R = 0.165$ ,  $P = 5.26e-04$ ), CD4+ T cells ( $R = 0.341$ ,  $P = 5.47e-14$ ), macrophages ( $R = 0.468$ ,  $P = 8.12e-26$ ), neutrophils ( $R = 0.468$ ,  $P = 2.56e-26$ ), and DCs ( $R = 0.408$ ,  $P = 1.16e-19$ ) indicated a positive association. In LIHC, PRKDC expression was positively correlated with B cells ( $R = 0.324$ ,  $P = 7.12e-10$ ), CD8+ T cells ( $R = 0.205$ ,  $P = 1.33e-04$ ), CD4+ T cells ( $R = 0.358$ ,  $P = 7.42e-12$ ), macrophages ( $R = 0.405$ ,  $P = 6.99e-15$ ), neutrophils ( $R = 0.422$ ,  $P = 2.38e-16$ ), and DCs ( $R = 0.42$ ,  $P = 5.73e-16$ ), but was not associated with tumor purity. About THCA, PRKDC gene shown a negative association with CD8+ T cells ( $R = -0.437$ ,  $P = 3.61e-24$ ), but there was a positive association with B cells ( $R = 0.672$ ,  $P = 1.20e-64$ ), CD4+ T cells ( $R = 0.665$ ,  $P = 1.01e-63$ ), macrophages ( $R = 0.715$ ,  $P = 1.09e-77$ ), neutrophils ( $R = 0.428$ ,  $P = 3.65e-23$ ), and DCs ( $R = 0.386$ ,  $P = 1.10e-18$ ), but was not related to tumor purity. We employed the SangerBox online tool to discuss the relevance between PRKDC and immune cell subtypes in the TME (Fig. 5B). The results indicated that CD56dim natural killer cell, memory B cell, monocyte and type 2 T helper cell had the strongest correlation with PRKDC expression.

### 3.6. Immune checkpoint genes, TMB, and MSI analysis results

Immune checkpoints are important immune system regulator. Some tumors can protect themselves from attack by taking advantage of immune checkpoint genes. We used SangerBox platform to investigate the association between PRKDC and the immune checkpoint genes across different tumors, as shown in Figure 6A. For example, in LIHC, PRKDC was positively correlated with expression of CD200, CD200R1, CD244, CD27, BTNL2, CD274, CD276, CD44, HAVCR4, CD28, CD80, CD86, CTLA4, HHLA2, ICOSLG, LGALS9, NRP1, TNFSF15, ICOS, LAG3, LAIR1, PDCD1, TNFSF4, TNFSF9, VSIR, VTCN1, TIGIT, TNFRSF8, and TNFSF18. TMB and MSI are important factors that influence tumor initiation and development, as well as tumor immunotherapy response. Higher TMB was related to better OS and better outcomes from immune checkpoint inhibitor (ICI) therapy in tumor patients.<sup>[27]</sup> Radar plots (Fig. 6B) showed a positive association between PRKDC and TMB in KICH ( $P = .0027$ ) and LUAD ( $P = 1.1e-05$ ), but a negative association in THCA ( $P = .008$ ) and COAD ( $P =$

$.021$ ). We also observed a positive association between PRKDC expression and MSI for GBM ( $P = .033$ ), OV ( $P = .0018$ ), LUSC ( $P = .0013$ ), SARC ( $P = .007$ ), and KIRC ( $P = .0048$ ) but a negative correlation for prostate adenocarcinoma ( $P = 1.1e-05$ ), SKCM ( $P = .009$ ), THCA ( $P = .00015$ ), HNSC ( $P = .0059$ ), and lymphoid neoplasm diffuse large B-cell lymphoma ( $P = .00044$ ) (Fig. 6C). In summary, all these data suggested a broad association between high PRKDC expression and tumor immu-

### 3.7. PPI network and enrichment analysis results

To investigate the mechanism of PRKDC in tumorigenesis, this work employed GeneMANIA online platform to build a PPI network for PRKDC. As it is shown in Figure 7A, PRKDC shared a same pathway with XRCC4, XRCC5, XRCC6, and LIG4, and PRKDC had strong physical interactions with XRCC5 and XRCC6. The aforementioned 2 genes encode Ku80 and Ku70, respectively, which make up a heterodimer protein critical for NHEJ pathway of DNA repair. As a regulatory subunit, the Ku80/Ku70 heterodimer can induce a 100-fold increase in affinity between DNA-PKcs and DNA. LIG4 is also required for NHEJ because it can form a complex with XRCC4 that further interacts with the DNA-PK.<sup>[28]</sup>

Enrichment analysis of high PRKDC expression and low PRKDC expression was performed using GSEA (Fig. 7B–E). The top few KEGG enrichment terms in the high PRKDC expression group named “cell cycle,” “oocyte meiosis,” and “RNA-degradation.” HALLMARK revealed enrichment terms of high PRKDC expression in “mitotic spindle,” “G2M check point,” and “P13K-AKT-MTOR signaling.”

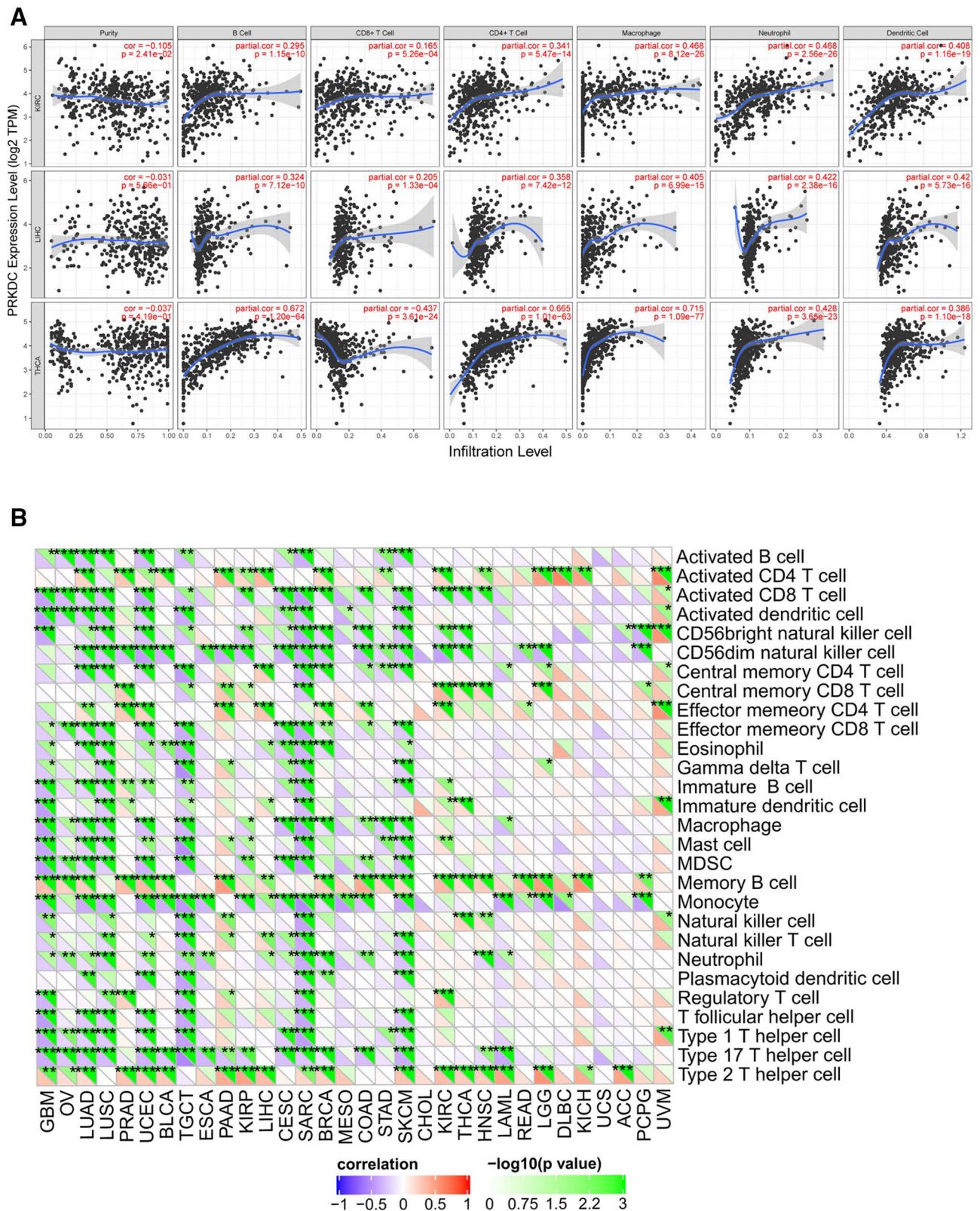
## 4. Discussion

Functional associations between PRKDC and various tumors have been reported in many publications.<sup>[10,29–33]</sup> It is unclear that whether PRKDC is involved in the etiology of various cancers via common molecular processes. There was no publication with PRKDC pan-cancer analysis from a holistic oncology approach, according to a literature search. Therefore, we conducted a comprehensive study of PRKDC genes in 33 types of cancers based on TCGA, GTEx, and CPTAC databases, including gene expression, gene alteration, protein phosphorylation, TME, and biological pathways.

Inducing irreversible DNA damage is considered to be a significant therapeutic strategy for cancer. DNA-dependent protein kinase complex (DNA-PK) consisted of DNA-PKcs and Ku80/Ku70 heterodimer protein, and it is very important to DSB repair response. Thus, it explains why DNA-PK is associated with a reduced response to DNA-damaging drugs as well as treatment resistance in various malignancies.<sup>[7,34,35]</sup> This study confirmed that PRKDC was substantially expressed in most tumors and was linked to a poor prognosis in a lot of tumor types.<sup>[36]</sup> The development of PRKDC targeted therapy has been prompted by these evidences and the fact that DNA-PKcs is a latent treatment target.<sup>[37,38]</sup> There are many pathways that can promote tumor cells survival and proliferation, so knowing PRKDC's regulatory mechanisms is crucial for designing effective therapies to inhibit PRKDC.

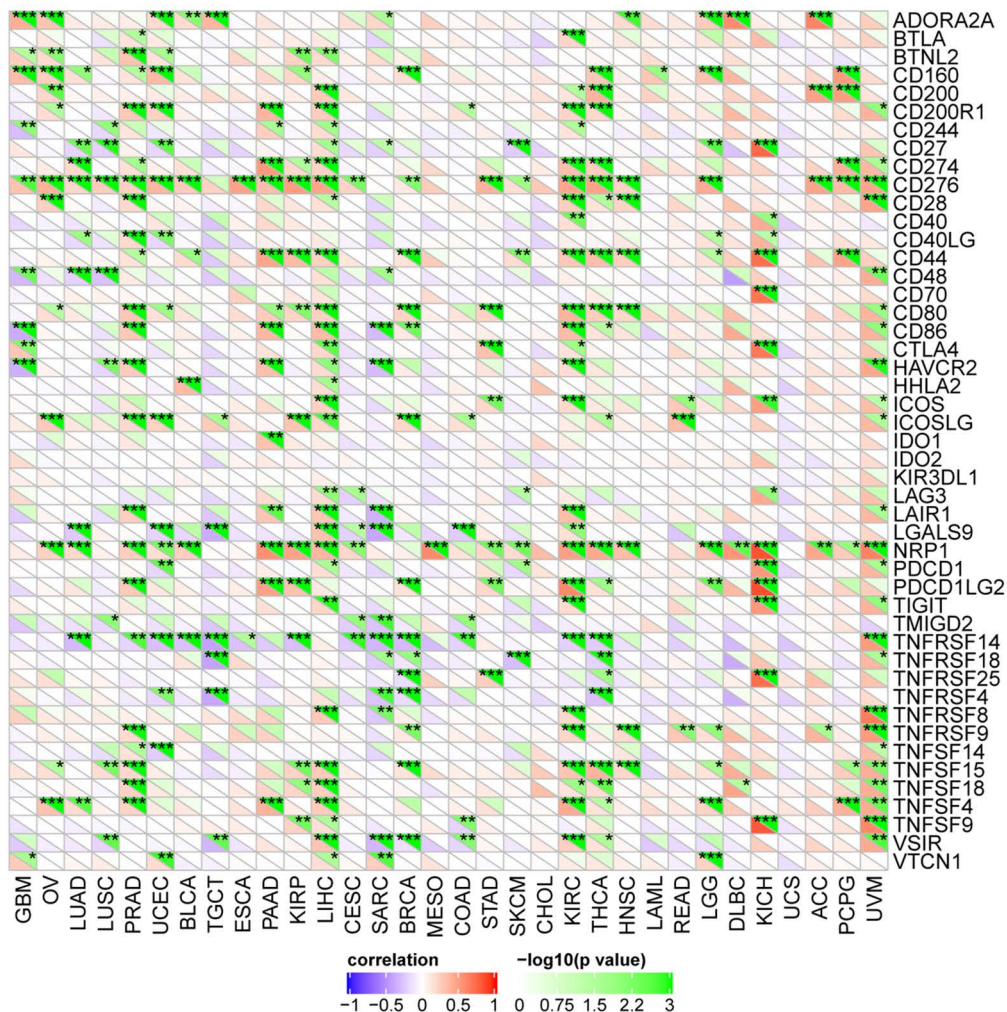
Only 10 cases have been recorded with disease-causing mutations in PRKDC. Six of them exhibited severe combined immunodeficiency, including granulomas and autoimmunity.<sup>[39–41]</sup> It results from that PRKDC mutation decreases affinity between DNA-PKcs and DNA and jeopardizes the activity of Artemis, which is necessary for V(D)J recombination as well. We found that over 18% of UCEC patients and about 12% of STAD patients presented alteration in PRKDC and the alteration correlates with better survival. In a recent study,<sup>[42]</sup> PRKDC mutation was found to be related to higher TMB, elevated mRNA levels of immunity-related genes as well as improved response



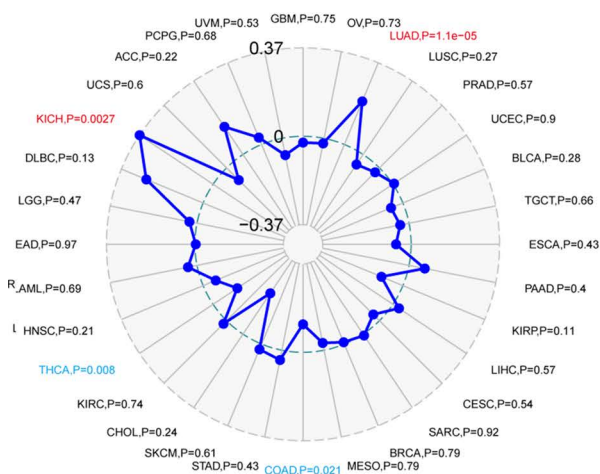


**Figure 5.** Coefficient of TIICs and PRKDC across TCGA tumors. (A) KIRC, LIHC, and THCA show the strongest association of PRKDC expression and immune infiltrating. (B) We identified and evaluated the relationship between PRKDC expression and 28 immune cell subtypes via SangerBox online tool. \* $P < .05$ , \*\* $P < .01$ , and \*\*\* $P < .001$ . KIRC = kidney renal clear cell carcinoma, LIHC = liver hepatocellular carcinoma, PRKDC = protein kinase, DNA-activated, catalytic subunit, TCGA = The Cancer Genome Atlas, THCA = thyroid carcinoma, TIICs = tumor-infiltrating immune cells.

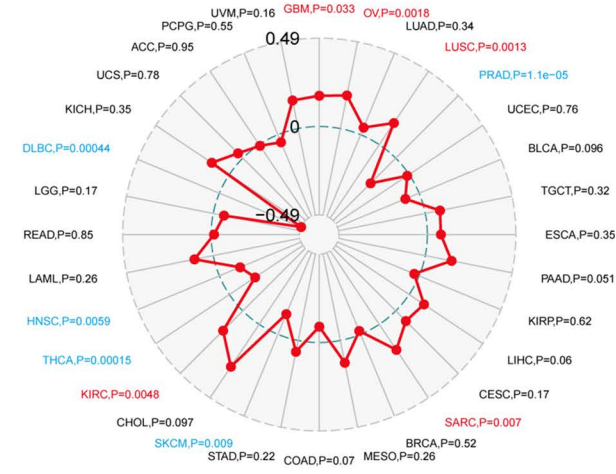
**A**



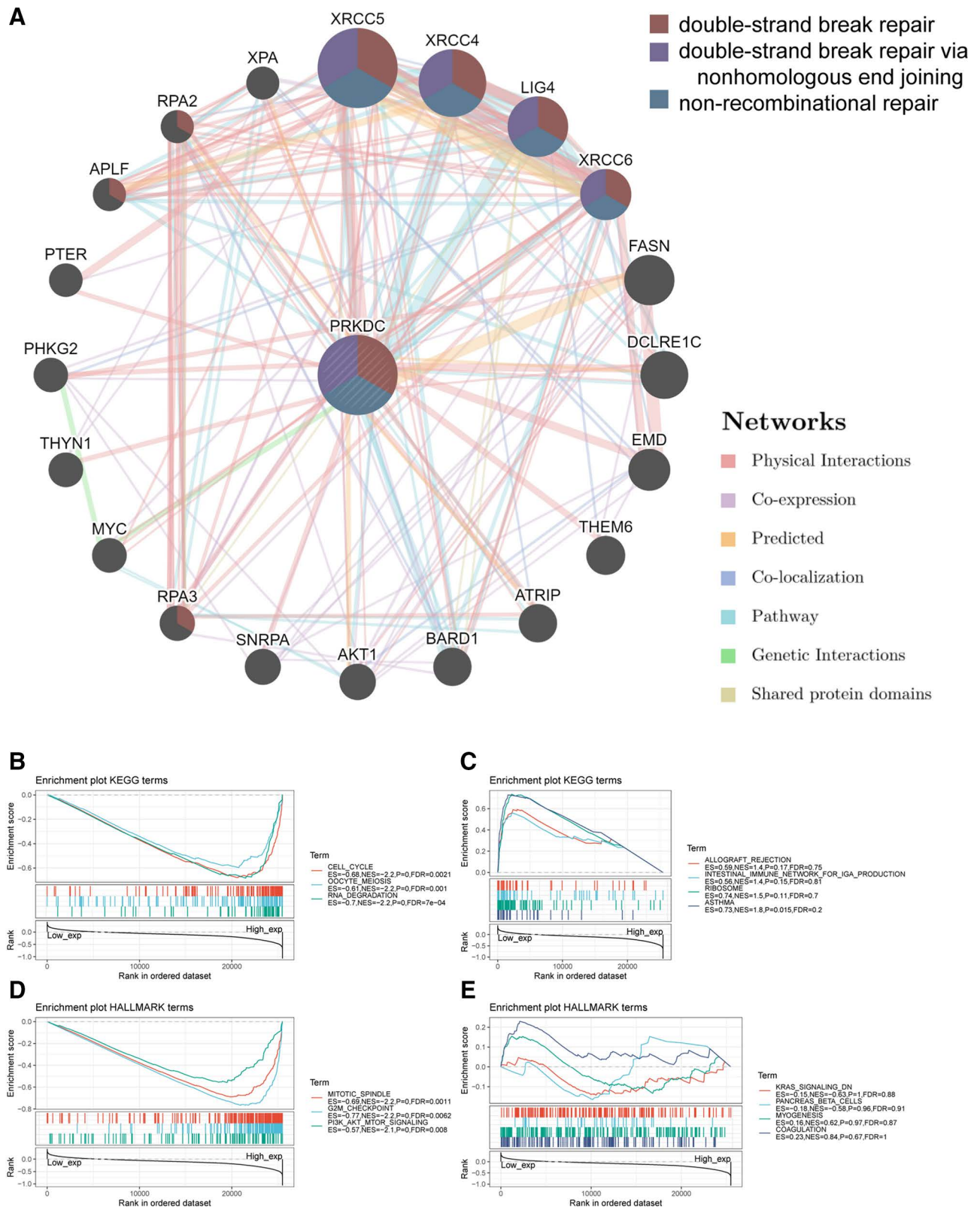
**B**



**C**



**Figure 6.** Correlation of PRKDC expression and checkpoint genes, TMB and MSI in tumors. (A) Relationship between PRKDC and 47 distinct immune checkpoint genes across different tumors. (B) The radar plot displays the association of the PRKDC gene and TMB. (C) The radar plot displays the association of the PRKDC gene and MSI. \* $P < .05$ , \*\* $P < .01$ , and \*\*\* $P < .001$ . MSI = microsatellite instability, PRKDC = protein kinase, DNA-activated, catalytic subunit, TMB = tumor mutation burden.



**Figure 7.** PRKDC-related genes enrichment analysis. (A) GeneMANIA online platform built a PPI network for PRKDC. Enriched gene sets by KEGG analysis in samples with high PRKDC expression (B) and low PRKDC expression (C). Enriched gene sets by HALLMARK analysis in samples with high PRKDC expression (D) and low PRKDC expression (E). Gene sets with  $|NES| > 1$ ,  $NOM P < .01$ , and  $FDR q < 0.25$  were defined as significant. The plot only showed a few of the most significant gene sets. KEGG = Kyoto encyclopedia of genes and genomes, PRKDC = protein kinase, DNA-activated, catalytic subunit, PPI = protein-protein interaction.

to ICI treatments in tumor individual. Another publication also discovered that PRKDC mutation was linked to an advanced TMB and MSI states in tumors, and knocked down PRKDC or used a DNA-PKcs inhibitor might improve ICI efficacy.<sup>[43]</sup> Above findings demonstrate that the PRKDC mutation might be used as a biomarker in immunotherapy.

We used the CPTAC dataset to compared the DNA-PKcs phosphorylation levels in the following 6 types of tumors: LUAD, colon cancer, clear cell RCC, breast cancer, UCEC, and ovarian cancer. It indicated a higher phosphorylation level at the S893, T2609, S2671, S3205, and S3995 locus in the primary tumors than normal controls. In vitro, DNA-PKcs experiences substantial autophosphorylation, causing it to dissociate from Ku-bound DNA and lose its kinase function.<sup>[44]</sup> In contrast, the capacity of DNA-PKcs to disassociate from Ku-DNA was diminished when the T2609 phosphorylation site was changed to alanine.<sup>[45,46]</sup> As a result, DNA-PKcs autophosphorylation was considered critical for the DNA repair process.

PRKDC expression has been linked to various levels of immune infiltration in tumors, particularly in KIRC, LIHC, and THCA, according to this study. There was a strong negative relevance between PRKDC and tumor purity in KIRC, showing that PRKDC is relatively abundant in the TME. However, the degree of PRKDC expression in LIHC and THCA was independent of tumor purity, indicating that it was expressed equally among tumor cell and the TME. For all 3 tumors, it showed a clearly positive associations between PRKDC expression and TIICs, but a negative association with CD8+ T cells in THCA. These distinctions indicated that there were variances among tumors in TME.

In the TME, immune system could identify and kill tumor cells. However, tumor cells can use a variety of tactics to survival and proliferate that inhibit the immune system.<sup>[47,48]</sup> Tumor immunotherapy, which includes of therapeutic antibodies, ICIs, tumor vaccines, and cell treatments, can help the body regain its natural antitumor immune response. We looked examined the link between PRKDC expression and 47 different immunological checkpoint genes, and found that high PRKDC expression may play a key role in immune evasion.

MSI was considered a biomarker for identifying individuals who may benefit from immunotherapies since it is linked to a higher cancer risk with important clinicopathological features, such as elevated TMB and more TIICs.<sup>[49]</sup> TMB might be adopted as a biomarker to forecast how well checkpoint blockades may work.<sup>[50,51]</sup> We showed evidence of a link between PRKDC expression and MSI or TMB across pan-cancers in this work, as well as the potential function of PRKDC in oncology immunological and as a predictive biomarker for a variety of malignancies.

In conclusion, our study revealed statistical associations between PRKDC expression and clinical outcome, genetic mutation, protein phosphorylation, immune cells infiltration, immune checkpoints, TMB, and MSI across pan-cancers, and confirmed PRKDC's latent involvement in tumor immunology and as a therapeutic target and prognostic biomarker. Unfortunately, this work merely performed a bioinformatics analysis of PRKDC across several databases and platforms, lacking of in vitro or in vivo experiments. Therefore, further cellular and molecular mechanistic research on PRKDC is desired to better understand the role of PRKDC in tumors.

### Author contributions

X.S. and X.Y. designed this work. F.Y. performed the bioinformatics analyses. All authors participated in writing the manuscript. All authors read and approved the final manuscript.

### Acknowledgments

The authors sincerely acknowledge the publicly available of the GTEx (<https://gtexportal.org/home/>) database and TCGA (<https://www.cancer.gov/>) database.

### References

- Hoelijmakers JH. Genome maintenance mechanisms for preventing cancer. *Nature*. 2001;411:366–74.
- Goodarzi AA, Jeggo PA. The repair and signaling responses to DNA double-strand breaks. *Adv Genet*. 2013;82:1–45.
- Davis AJ, Chen DJ. DNA double strand break repair via non-homologous end-joining. *Transl Cancer Res*. 2013;2:130–43.
- Hartley KO, Gell D, Smith GC, et al. DNA-dependent protein kinase catalytic subunit: a relative of phosphatidylinositol 3-kinase and the ataxia telangiectasia gene product. *Cell*. 1995;82:849–56.
- Goodwin JF, Knudsen KE. Beyond DNA repair: DNA-PK function in cancer. *Cancer Discov*. 2014;4:1126–39.
- Zhang Y, Wen GM, Wu CA, et al. PRKDC is a prognostic marker for poor survival in gastric cancer patients and regulates DNA damage response. *Pathol Res Pract*. 2019;215:152509.
- Shintani S, Mihara M, Li C, et al. Up-regulation of DNA-dependent protein kinase correlates with radiation resistance in oral squamous cell carcinoma. *Cancer Sci*. 2003;94:894–900.
- Lee SW, Cho KJ, Park JH, et al. Expressions of Ku70 and DNA-PKcs as prognostic indicators of local control in nasopharyngeal carcinoma. *Int J Radiat Oncol Biol Phys*. 2005;62:1451–7.
- Ihara M, Ashizawa K, Shichijo K, et al. Expression of the DNA-dependent protein kinase catalytic subunit is associated with the radiosensitivity of human thyroid cancer cell lines. *J Radiat Res*. 2019;60:171–7.
- Cornell L, Munck JM, Alsinet C, et al. DNA-PK-A candidate driver of hepatocarcinogenesis and tissue biomarker that predicts response to treatment and survival. *Clin Cancer Res* 2015;21:925–33.
- Evert M, Frau M, Tomasi ML, et al. Deregulation of DNA-dependent protein kinase catalytic subunit contributes to human hepatocarcinogenesis development and has a putative prognostic value. *Br J Cancer*. 2013;109:2654–64.
- Doherty RE, Bryant HE, Valluru MK, et al. Increased non-homologous end joining Makes DNA-PK a promising target for therapeutic intervention in uveal melanoma. *Cancers*. 2019;11:1278.
- Davidson D, Amrein L, Panasci L, et al. Small molecules, inhibitors of DNA-PK, targeting DNA repair, and beyond. *Front Pharmacol*. 2013;4:5.
- Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, et al. TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res*. 2020;48:W509–w14.
- Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res*. 2017;45:W98–W102.
- Chandrashekar DS, Bashel B, Balasubramanya SAH, et al. UALCAN: a portal for facilitating tumor subgroup gene expression and survival analyses. *Neoplasia (New York, NY)*. 2017;19:649–58.
- Rudnick PA, Markey SP, Roth J, et al. A Description of the clinical proteomic tumor analysis Consortium (CPTAC) common data analysis pipeline. *J Proteome Res*. 2016;15:1023–32.
- Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signaling*. 2013;6:11.
- Li T, Fan J, Wang B, et al. TIMER: a web server for comprehensive analysis of tumor-infiltrating immune cells. *Cancer Res*. 2017;77:e108–10.
- Yoshihara K, Shahmoradgoli M, Martínez E, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun*. 2013;4:2612.
- Sha D, Jin Z, Budczies J, et al. Tumor mutational burden as a predictive biomarker in solid tumors. *Cancer discovery* 2020;10:1808–25.
- Hause RJ, Pritchard CC, Shendure J, et al. Classification and characterization of microsatellite instability across 18 cancer types. *Nat Med*. 2016;22:1342–50.
- Mostafavi S, Ray D, Warde-Farley D, et al. GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. *Genome Biol*. 2008;9(Suppl 1):S4.
- Subramanian A, Tamayo P, Mootha VK, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci USA*. 2005;102:15545–50.
- Dephoure N, Zhou C, Villén J, et al. A quantitative atlas of mitotic phosphorylation. *Proc Natl Acad Sci USA*. 2008;105:10762–7.
- Fridman WH, Galon J, Dieu-Nosjean MC, et al. Immune infiltration in human cancer: prognostic significance and disease control. *Curr Top Microbiol Immunol*. 2011;344:1–24.
- Samstein RM, Lee CH, Shoushtari AN, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet*. 2019;51:202–6.

- [28] Andres SN, Vergnes A, Ristic D, et al. A human XRCC4-XLF complex bridges DNA. *Nucleic Acids Res.* 2012;40:1868–78.
- [29] Noguchi T, Shibata T, Fumoto S, et al. DNA-PKcs expression in esophageal cancer as a predictor for chemoradiation therapeutic sensitivity. *Ann Surg Oncol.* 2002;9:1017–22.
- [30] Lee HS, Choe G, Park KU, et al. Altered expression of DNA-dependent protein kinase catalytic subunit (DNA-PKcs) during gastric carcinogenesis and its clinical implications on gastric cancer. *Int J Oncol.* 2007;31:859–66.
- [31] Zhang Y, Yang WK, Wen GM, et al. High expression of PRKDC promotes breast cancer cell growth via p38 MAPK signaling and is associated with poor survival. *Mol Genet Genomic Med.* 2019;7:e908.
- [32] Qu S, Zhu XD, Liang SX, et al. Expression and clinical significance of DNA-PKcs in nasopharyngeal carcinoma tissues. *Chin J Cancer Prev Treat.* 2009;16:916–8.
- [33] Lü Y, Zhang H, Li Y, et al. dependent protein kinase catalytic subunit and P16 in colorectal carcinoma]. *Zhonghua yi xue za zhi.* 2008;88:2025–9.
- [34] Beskow C, Skikuniene J, Holgersson A, et al. Radioresistant cervical cancer shows upregulation of the NHEJ proteins DNA-PKcs, Ku70 and Ku86. *Br J Cancer.* 2009;101:816–21.
- [35] Um JH, Kwon JK, Kang CD, et al. Relationship between antiapoptotic molecules and metastatic potency and the involvement of DNA-dependent protein kinase in the chemosensitization of metastatic human cancer cells by epidermal growth factor receptor blockade. *J Pharmacol Exp Ther.* 2004;311:1062–70.
- [36] Hosoi Y, Watanabe T, Nakagawa K, et al. Up-regulation of DNA-dependent protein kinase activity and Sp1 in colorectal cancer. *Int J Oncol.* 2004;25:461–8.
- [37] Kurimasa A, Kumano S, Boubnov NV, et al. Requirement for the kinase activity of human DNA-dependent protein kinase catalytic subunit in DNA strand break rejoining. *Mol Cell Biol.* 1999;19:3877–84.
- [38] Zhao Y, Thomas HD, Batey MA, et al. Preclinical evaluation of a potent novel DNA-dependent protein kinase inhibitor NU7441. *Cancer Res.* 2006;66:5354–62.
- [39] Esenboga S, Akal C, Karaatmaca B, et al. Two siblings with PRKDC defect who presented with cutaneous granulomas and review of the literature. *Clin Immunol.* 2018;197:1–5.
- [40] Mathieu AL, Verronese E, Rice GI, et al. PRKDC mutations associated with immunodeficiency, granuloma, and autoimmune regulator-dependent autoimmunity. *J Allergy Clin Immunol.* 2015;135:1578–88.e5.
- [41] van der Burg M, Ijspeert H, Verkaik NS, et al. A DNA-PKcs mutation in a radiosensitive T-B- SCID patient inhibits artemis activation and nonhomologous end-joining. *J Clin Invest.* 2009;119:91–8.
- [42] Chen Y, Li Y, Guan Y, et al. Prevalence of PRKDC mutations and association with response to immune checkpoint inhibitors in solid tumors. *Mol Oncol.* 2020;14:2096–110.
- [43] Tan KT, Yeh CN, Chang YC, et al. PRKDC: new biomarker and drug target for checkpoint blockade immunotherapy. *J ImmunoTher Cancer.* 2020;8:e000485.
- [44] Chan DW, Lees-Miller SP. The DNA-dependent protein kinase is inactivated by autophosphorylation of the catalytic subunit. *J Biol Chem.* 1996;271:8936–41.
- [45] Jette N, Lees-Miller SP. The DNA-dependent protein kinase: a multifunctional protein kinase with roles in DNA double strand break repair and mitosis. *Progress in Biophysics and Molecular Biology.* 2015;117:194–205.
- [46] Hammel M, Yu Y, Mahaney BL, et al. Ku and DNA-dependent protein kinase dynamic conformations and assembly regulate DNA binding and the initial non-homologous end joining complex. *J Biol Chem.* 2010;285:1414–23.
- [47] Colegio OR, Chu NQ, Szabo AL, et al. Functional polarization of tumour-associated macrophages by tumour-derived lactic acid. *Nature.* 2014;513:559–63.
- [48] Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity.* 2013;39:782–95.
- [49] Sahin IH, Akce M, Alese O, et al. Immune checkpoint inhibitors for the treatment of MSI-H/MMR-D colorectal cancer and a perspective on resistance mechanisms. *Br J Cancer.* 2019;121:809–18.
- [50] Liu L, Bai X, Wang J, et al. Combination of TMB and CNA stratifies prognostic and predictive responses to immunotherapy across metastatic cancer. *Clin Cancer Res.* 2019;25:7413–23.
- [51] Chan TA, Yarchoan M, Jaffee E, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann Oncol.* 2019;30:44–56.