Cell Division Cycle-Associated Protein 8 in Pan-Cancer: A Comprehensive Prognostic and Immune Infiltration Investigation

Wanrong Zheng
Minxi Vocational and Technical College

Chengqian Zhong
Longyan First Affiliated Hospital of Fujian Medical University Longyan

Fobao Lai (lfbjxgz@126.com)
Longyan First Affiliated Hospital of Fujian Medical University Longyan

Research Article

Keywords: CDCA8, prognosis, pan-cancer, Immune Infiltration, cell cycle

Posted Date: October 11th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-3299462/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License

Additional Declarations: No competing interests reported.
Abstract

The Cell Division Cycle-Associated Protein 8 (CDCA8) protein is a vital player in the process of cell division and can influence numerous physiological and pathological events in the body by triggering certain proteins such as cell cycle-controlling proteins, transcription factors, and signal transmission molecules. Despite numerous studies indicating that dysregulation of CDCA8 is prevalent in human cancers, no systematic pan-cancer analysis has been conducted. In this study, we used The Cancer Genome Atlas (TCGA), Human Protein Atlas (HPA), and Gene Expression Omnibus (GEO) datasets, and several bioinformatics tools to investigate the role of CDCA8 in 33 different tumor types. The results showed that Patients with malignancies that overexpressed CDCA8, such as adrenocortical carcinoma, kidney renal clear cell carcinoma, and liver hepatocellular carcinoma, typically had poor overall survival (OS). We also found that CDCA8 expression was favorably correlated with immune cell infiltration levels in numerous human cancer types. Furthermore, GSEA results showed that overexpression of CDCA8 in human malignancies may accelerate the development of cancer by triggering a number of oncogenic signaling pathways. In conclusion, our comprehensive assessment of the oncogenic activity of CDCA8 in several human cancer types suggests that CDCA8 can be used as a potential therapeutic target and prognostic biomarker in various human cancer types.

Introduction

The importance of pan-cancer research is the use of cross-cancer similarities to diagnose and cure more malignancies. This makes it useful for cancer detection and treatment to distinguish cancer markers among the various cancer types.

It is crucial to understand how abnormal cell cycle regulation affects the development and progression of cancer because it may result in excessive cell proliferation and cancer promotion[1, 2]. Therefore, understanding the elements that affect cell cycle regulation is crucial for understanding how cancer develops and progresses. Various phases of the cell cycle require the expression of genes that control cell division. The connection between genes associated with cell cycle control and carcinogenesis has also been reported in previous studies, and cell division cycle-associated protein 8 (CDCA8) has been discovered to be highly expressed in ovarian, liver, and prostate cancers[3–5].

A study that offers a comprehensive pan-cancer examination of CDCA8 is still lacking, despite the fact that increasing amounts of literature claim that CDCA8 plays an important biological role in the development of cancer. Therefore, we performed a pan-cancer investigation of CDCA8 in terms of gene expression, prognostic relevance, immunological association, and gene enrichment analysis using TCGA, GTEx, GEPIA, STRING, TIMER, Metascape, and other databases. This study aimed to advance the knowledge of CDCA8 in pan-cancer for cancer researchers.

Materials and methods
Gene Expression Analysis

To compare the expression profiles of CDCA8 in several tumor types and nearby normal tissues, researchers used TIMER2 (Tumor Immune Estimation Resource, version 2, http://timer.cistrome.org/). The GEPIA2 (Gene Expression Profiling Interactive Analysis, version 2) program (http://gepia2.cancer-pku.cn/#analysis) was used to obtain box plots of the Genotype-Tissue Expression (GTEx) database in cases where tumors often lacked or only had a small amount of normal tissue. All TCGA expression of CDCA8 in all TCGA tumors at various pathological phases was examined using the HEPIA2 tool. We examined the CPTAC (Clinical Proteomic Tumor Analysis Collaboration) dataset for protein expression[6]. The levels of expression of CDCA8 total protein or phosphoprotein have been studied in primary and normal tissues.

Immunohistochemistry (IHC) Staining

IHC images of CDCA8 protein expression in normal tissues and six tumor tissues, including BRCA, COAD, PAAD, LIHC, and CESC, were downloaded from the HPA (Human Protein Atlas) (http://www.proteinatlas.org/) and processed to assess variations in CDCA8 expression at the protein level.

Survival Prognosis Analysis

We obtained survival plots of CDCA8 across all TCGA cancers and the Overall Survival (OS) significance map data using GEPIA2. The expression thresholds utilized to separate the cohorts with high and low expression levels were cutoff high (50%) and cutoff low (50%) values[7]. The hypotheses were tested using log-rank tests.

Genetic Alteration Analysis

Data on alteration frequency, mutation type, altered site, and copy number alteration (CNA) across all TCGA tumors were gathered using the cBioPortal program (https://www.cbioportal.org/).

Immune Infiltiration Analysis

To examine the connection between CDCA8 expression and immune infiltrates in all TCGA tumors, the TIMER2 tool was employed. Several cells were chosen for in-depth research, including cancer-associated fibroblasts, myeloid-derived suppressor cells, CD4 + T cells, and macrophage cells. The following algorithms were used for the estimations: TIMER, TIDE, CIBERSORT, CIBERSORT-ABS, QUANTISEQ, XCELL, MCPCOUNTER, and EPIC.

CDCA8-Related Gene Enrichment Analysis

The following key parameters were used to build a Homo Sapiens CDCA8 co-expression network with the STRING tool (version:11.0b) (https://string-db.org/), which is one of the active interaction sources. Evidence is the meaning of network edges. The maximum number of participants was 50. The minimum required interaction score was low confidence (0.150).
The 100 CDCA8-correlated genes with the most resemblance to CDCA8 expression pattern were obtained from TCGA datasets using the GEPIA2 "Similar Gene Detection" module. Using the gene symbols of these 100 genes as input gene symbols, the "clusterProfiler" R package (version:3.13) was used to do a Gene Ontology pathway enrichment study. The GEPIA2 "Correlation Analysis" module was also used to do pairwise gene correlation analysis.

**CDCA8-Protein Interaction Analysis**

An interaction network between the CDCA8 protein was built using the "Network" module of BioGRID (version:4.3) ([https://thebiogrid.org/](https://thebiogrid.org/)), with the layout set to "Concentric Circles."

**RESULTS**

**Transcriptional Level Analysis of CDCA8 in Pan-Cancer**

We first used the gene expression data from TCGA and GTEx databases to analyze the differential expression of CDCA8 in cancer and non-cancer tissues and created matching violin plots to investigate the expression of CDCA8 in pan-cancer. According to these findings, CDCA8 was highly expressed in the majority of tumors (Fig. 1A), which was roughly in line with the findings from the TCGA and GTEx databases (Fig. 1B). The expression of CDCA8 in 23 various tumor types with paired samples was also investigated using TCGA (Fig. 1C). Overall, most cancer types have substantial overexpression of CDCA8.

We also examined IHC results from the HPA database and compared them with TCGA data on CDCA8 expression. The study of the data in these two databases produced results that agreed with each other. Normal breast, colon, pancreatic, liver, and cervical tissues stained negatively or moderately with CDCA8, whereas malignant tissues stained strongly or moderately with CDCA8 (Figs. 2A–E).

**Survival Analysis Data**

The correlation between CDCA8 expression and prognosis and overall survival will be the next area of research. To explore the relationship between CDCA8 expression and prognosis in various tumor patients, cancer cases were divided into high- and low-expression groups based on the degree of CDCA8 expression. In malignancies including ACC (P = 3.9e-07), KIRC (P = 0.0059), LGG (P = 6.4e-09), and LIHC (P = 0.00026), high expression of CDCA8 was linked to a poor prognosis for Overall Survival (OS) (Fig. 3).

We also used the GEPIA2 tool to examine the relationship between CDCA8 expression and tumor pathological staging, and the results showed that for a few tumor types, including KIRC, ACC, BRCA, KICH, KIRP, LIHC, LUAD, PAAD, and THCA (Fig. 4, all P < P0.05), there were stage-specific expressional changes in CDCA8 expression.

**Genetic Alteration Analysis Data**

Genetic changes occur over time and can cause cancer in humans. Therefore, we investigated genetic changes in CDCA8 in human tumor samples. Our data shows that OV has the highest frequency of
CDCA8 modification (> 7%), with "amplification" as the main kind. With a frequency of ~ 2%, pheochromocytoma and paraganglioma had the highest incidence of "deep deletion" type CNA. We displayed additional mutations and where they are located within CDCA8 (Fig. 5B). No specific type of genetic modification was discovered, and some were located in the Borealin domain.

**Protein Phosphorylation Analysis Data**

It is well recognized that a crucial step in oncogenesis is the phosphorylation-dephosphorylation cascade. Serine and threonine residues in CDCA8 are phosphorylated as part of the post-translational modifications that are being made to it, followed by evaluation of CDCA8 phosphorylation in normal and primary tumor tissues. Using the CPTAC dataset, we conducted a more thorough analysis of three tumor types (breast cancer, lung adenocarcinoma, and head and neck squamous cell carcinoma). We discovered that breast cancer primary tumor tissues and lung adenocarcinoma primary tumor tissues had significantly higher levels of CDCA8 phosphorylation at S219, as shown in Fig. 6A and B. For head and neck squamous cell carcinoma, we discovered increased T204, S215, and S219 phosphorylation levels (Figs. 6C).

**Immune Infiltration Analysis Data**

We applied TIMER, CIBERSORT, CIBERSORT-ABS, TIDE, XCELL, MCPCounter, QuantiSeq, and EPIC algorithms to examine the relationship between CDCA8 expression and immune cell infiltration in various TCGA tumor types. Our goal was to determine whether there is a relationship between CDCA8 expression and immune cell infiltration in pan-cancer. Interestingly, we found a favorable association between CDCA8 expression and the projected MDSC and Th2 cell infiltration values for different cancer types. A positive link was established between CDCA8 expression and M1 macrophage infiltration in various cancer types; however, a negative correlation was observed between CD4 + effector memory T cells and CDCA8 expression in various cancer types (Fig. 7).

Previous research has revealed that various immune cells that infiltrate tumors are regulated by cancer-associated fibroblasts in the stroma (Chen and Song, 2019). Therefore, we used EPIC and MCPCounter algorithms to examine the relationship between CDCA8 expression and cancer-associated fibroblast infiltration in various malignancies. In ACC, CESC, KIRC, KIRP, LGG, LIHC, LUAD, MESO, PCPG, SKCM, and thyroid carcinoma (THCA), CDCA8 expression was positively correlated with cancer-associated fibroblast infiltration (Fig. 8).

**CDCA8-Related Gene Enrichment Analysis**

We employed GEPIA2 to extract the top 100 genes from all tumor types in TCGA datasets with expression patterns comparable to those of CDCA8 to examine the functional mechanism of CDCA8 in carcinogenesis (Supplementary Table S1). These genes had a high degree of association with the regulation of the cell cycle or mitosis, according to a gene ontology enrichment analysis (Fig. 9A). The STRING tool was used to gather 50 genes that were co-expressed with CDCA8 to validate the findings of the Gene Ontology enrichment study. These 50 genes had strong inter-gene interactions, as shown in
Fig. 9B. The genes were also enriched in cell cycle and mitotic control (Supplementary Table S2). These results prompted us to investigate whether CDCA8 interacts with important proteins involved in the cell cycle and mitotic control to influence these biological processes. The BioGRID4.3 database states that CDCA8, PARP1, AURKB, ANLN, KIF23, H3F3A, EGFR, HIST1H3A, KRAS and CBX5 physically interact with one another (Fig. 9C), which have well-characterized functions in the cell cycle, mitotic regulation, and tumorigenesis (Supplementary Figure S3) (Yuan and Chen, 2013; Laurini et al., 2020). Additionally, PARP1, AURKB, and KIF23 expression levels were highly correlated with CDCA8 expression (Figs. 9D, E). These findings led us to hypothesize that CDCA8 may contribute to tumorigenesis in malignancies by regulating the cell cycle and increasing cell division.

Discussion

An increasing number of studies have examined the role of CDCA8 in various illnesses, including cancer. However, it is unclear whether CDCA8 contributes to the more widespread pathways that are responsible for tumor pathogenesis or whether it plays a role in the oncogenesis of specific tumor types. Here, we performed CDCA8 pan-cancer analysis. Based on TCGA data, we examined the expression of CDCA8 in 33 distinct cancers. Using the CPTAC and GEO databases, we methodically gathered and integrated information on proteins, phospho-proteins, other molecular characteristics, and genetic changes.

In this study, we initially used TCGA database to investigate the mRNA expression level of CDCA8 in cancer tissues and normal tissues to acquire a thorough understanding of the differential expression of CDCA8 in pan-cancer. There are various high expression levels in more than ten different forms of cancer. However, while searching the TCGA database, we discovered that there were very few sequencing results for normal or paracancerous tissues. As a result, many cancer samples lack the corresponding transcriptomes for normal or paracancerous tissues such as ACC, DLBC, LAML, LGG, MESO, OV, TGCT, and UCS. To obtain more complete transcriptome data, we integrated TCGA database with the GTEx database, which offers additional information about normal tissue expression. The mixed findings of TCGA and GTEx databases indicated that practically all human malignancies have abnormally elevated CDCA8 mRNA expression. Next, we used the GEPIA database to examine the possible predictive significance of CDCA8 in pan-cancers. According to our overall survival analysis, CDCA8 overexpression may serve as a biomarker for prognosis in a number of malignancies, including ACC, KIRC, KIRP, LGG, LIHC, LUAD, MESO, PAAD, SARC, and SKCM. Different levels of elevated expression of CDCA8 are present in several cancer types compared to equivalent normal samples. Compared to patients with low CDCA8 expression, those with these cancer types in the high CDCA8 expression group had a worse prognosis.

Immune cells extensively interact with cancer cells and exert essential effects on cancer migration and metastasis in various tumor types[8, 9]. Recent studies have also reported that the tumor immune microenvironment is associated with the expression of various genes[10]. Fibroblasts are activated inside the tumor microenvironment by a variety of inflammatory cytokines released by cancer, host immune, and stromal cells. Cancer-associated fibroblasts (CAFs) are the name given to these activated fibroblasts. CAFs have a significant impact on neighboring cells through the production of soluble substances,
including cytokines and chemokines, as well as the ECM. By encouraging cancer cell development, boosting pro-tumor immune responses, altering the ECM, affecting tumor cell treatment resistance, and encouraging angiogenesis, they aid in the progression and spread of tumors[11, 12]. In this study, we found that CDCA8 expression was positively correlated with CAFs infiltration in several tumor types. We also found a favorable association between CDCA8 expression and projected MDSC and Th2 cell infiltration values for different cancer types. MDSCs, which accumulate in patients who do not respond well to cancer immunotherapies, are a heterogeneous population of immature myeloid cells with various potent immunosuppressive activities involving different immunocompetent cells[13, 14].

Using STRING and GEPIA2, we identified several genes that were co-expressed with CDCA8 across different tumors and other tissues. Gene enrichment analysis revealed that these genes were strongly correlated with the cell cycle or mitosis regulation, which is consistent with the results of previous studies.

Additionally, there are still certain limitations to our study, such as the lack of clinical and laboratory data. Therefore, in the future, we will investigate the biological significance of CDCA8 and its prognostic value in clinicopathological samples and cancer cell assays.

In conclusion, CDCA8 is frequently overexpressed in a variety of malignancies, and both its expression and genetic changes are statistically related to the clinical outcomes of individuals with particular tumors. Additionally, analyses of immune infiltration and CDCA8-related gene enrichment provided plausible pathways by which CDCA8 may control the cancer cell cycle, DNA repair, and tumor immunity. To better more the CDCA8’s potential use in cancer therapy and prognosis prediction, additional experimental and clinical investigations are necessary.

**Declarations**

- **Funding**
  No funding.

- **Competing interests**
  The authors declare that they have no conflict of interest.

- **Availability of data and materials**
  The original TCGA data that support the findings of our study are available in the NCI GDC Data portal repository at https://portal.gdc.cancer.gov/repository. Gene Expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/) data were obtained from NCBI GEO.

- **Ethics approval and consent to participate**
  Ethnical is not applicable because these data are from public database.
• Consent for publication

Not applicable.

• Author Contributions

Wanrong Zheng analyzed the data and wrote the manuscript. Fobao Lai has reviewed the manuscript. All authors have contributed to the manuscript and approved the submitted version.

References


Supplementary Information

Supplementary Figures and Supplementary Tables are not available with this version.

Figures
Expression distribution of CDCA8 in pan-cancer. The mRNA expression of CDCA8 in pan-cancer. (A) The mRNA expression of CDCA8 in 33 tumors in TCGA_GTEx samples. (B) The mRNA expression of CDCA8 in 33 tumors in TCGA database. (C) Expression of CDCA8 in paired samples of 18 tumors in TCGA database. ACC, adrenocortical carcinoma; BLCA, bladder urothelial carcinoma; BRCA, breast invasive carcinoma; CESC, cervical and endocervical cancers; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; DLBC, lymphoid neoplasm diffuse large B-cell lymphoma; ESCA, esophageal carcinoma; GBM, glioblastoma multiforme; HNSC, head and neck squamous cell carcinoma; 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; OV, ovarian serous cystadenocarcinoma; PAAD, pancreatic adenocarcinoma; PCPG, pheochromocytoma and paraganglioma; PRAD, prostate adenocarcinoma; READ, rectum adenocarcinoma; SARC, sarcoma; SKCM, skin cutaneous melanoma; STAD, stomach adenocarcinoma; STES, stomach and esophageal carcinoma; TGCT, testicular germ cell tumors; THCA, thyroid carcinoma; THYM, thymoma; UCEC, uterine corpus endometrial carcinoma; UCS, uterine carcinosarcoma; UVM, uveal melanoma. (ns, $P > 0.05$; *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$).
Figure 2

Comparison of CDCA8 gene expression between normal and tumor tissues (left) and immunohistochemistry images in normal (middle) and tumor (right) tissues. CDCA8 protein expression was significantly higher in BRCA, COAD, PAAD, LIHC and CESC. (A) Breast. (B) Colon. (C) Pancreas. (D) Liver. (E) Cervical. **P < 0.01; ***P < 0.001.
**Figure 3**

Correlation between CDCA8 expression and disease-free survival in patients with different TCGA tumor types. GEPIA2 was used to build a survival map (A) and conduct disease-free survival (B) analyses. The survival map and Kaplan-Meier plots with significant results are displayed. The 95% confidence intervals of disease-free survival are indicated by red and blue dotted lines for high and low CDCA8 groups, respectively.
Figure 4

Stage-dependent expression of CDCA8. The main pathological stages (stages I, II, III, and IV) of KIRC, ACC, BRCA, KICH, KIRP, LIHC, LUAD, PAAD, and THCA were assessed and compared using TCGA data. The expression levels are shown as Log2 (TPM+1).
Figure 5
Mutation status of CDCA8 in TCGA tumors. Mutation status of CDCA8 in TCGA tumors was analyzed using the cBioPortal tool. The alteration frequency with mutation type (A) and mutation site (B) are displayed.
Figure 6

Tumor-associated protein phosphorylation of CDCA8. Comparison of the level of CDCA8 phosphoproteins (S215, S219 and T204 sites) between normal tissue and primary tissue of selected tumors. (A–C) Box plot representation of CDCA8 phosphoprotein levels in breast cancer, lung adenocarcinoma and head and neck squamous cell carcinoma.
Figure 7

The correlation between CDCA8 expression level and infiltration of myeloid-derived suppressor cells, CD4+ T cells and macrophage cells. TIMER, CIBERSORT, CIBERSORT-ABS, TIDE, XCELL, MCPCounter, QUANTISEQ and EPIC algorithms were used for the correlative analysis of the level of myeloid-derived suppressor cells (A), CD4+ T cells (B) and macrophage cells (C) and the expression levels of the CDCA8 gene across all tumors in TCGA. The red color indicates a positive correlation (0-1), while the blue color represents a negative correlation (~1-0). The correlation with P-value <0.05 is considered as statistically significant. Statistically non-significant correlations values are marked with a cross.
Figure 8

Correlation between CDCA8 expression and cancer-associated fibroblast immune infiltration. EPIC and MCPCounter algorithms were used to calculate the correlation between CDCA8 expression and cancer-associated fibroblast immune infiltration in all tumor types from TCGA.
Figure 9

CDCA8-related gene enrichment analysis. (A) Gene Ontology (GO) analysis and KEGG analysis of the top 100 genes co-expressed with CDCA8 obtained by the GEPIA2. (B) Co-expression network of 50 genes co-expressed with CDCA8 obtained by the STRING tool. (C) CDCA8-protein interactions obtained by BioGRID. (D, E) Correlation analysis between CDCA8 and PARP1, AURKB and KIF23 conducted by GEPIA2 across all tumor samples from TCGA.