doi: 10.4149/gpb_2021033

Identification of hub genes in chronic pancreatitis and analysis of association with pancreatic cancer *via* bioinformatic analysis

Hongxuan Li, Chenjun Hao, Qiu Yang, Wenqi Gao, Biao Ma and Dongbo Xue

Department of General Surgery, Key Laboratory of Hepatosplenic Surgery, Ministry of Education, The First Affiliated Hospital of Harbin Medical University, Harbin, China

Abstract. Chronic pancreatitis (CP), a fibroinflammatory disease, is a potential risk factor for pancreatic cancer. This study attempted to identify and analyze the key genes involved in CP development and their association with pancreatic cancer. The GSE41418 dataset was obtained from the Gene Expression Omnibus database. Gene ontology and Kyoto Encyclopedia of Genes and Genomes pathway analyses were performed on common differentially expressed genes. A protein-protein interaction network was constructed by using the STRING database. The expression and prognostic value of hub genes in pancreatic cancer were analyzed by Gene Expression Profiling Interactive Analysis (GEPIA) and UALCAN databases. The results showed that the upregulated genes primarily focused on the cell cycle, DNA replication, and phagosome activity. The PPI network was composed of 184 nodes and 925 edges. The 10 hub genes were screened by CytoHubba, of which CCNB2, CDC6, CDK1 and CKS2 were observed to be differentially expressed in pancreatic cancer with CP, and all of them were detrimental to overall survival and recurrence-free survival of pancreatic cancer. In this study, we employed bioinformatic analysis to determine that CCNB2, CDC6, CDK1 and CKS2 may be key genes in the development of CP and pancreatic cancer.

Key words: Chronic pancreatitis — Pancreatic cancer — GEPIA — UALCAN — Prognostic value

Abbreviations: CCNB2, cyclin B2; CDC6, cell division cycle 6; CDK1, cyclin-dependent kinase 1; CKS2, CDC28 protein kinase regulatory subunit 2; CP, chronic pancreatitis; DEGs, differentially expressed genes; EMT, epithelial to mesenchymal transformation; FDR, false discovery rate; GEO, Gene Expression Omnibus; GEPIA, Gene Expression Profiling Interactive Analysis; GO, Gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; MCODE, molecular complex detection OS, overall survival; PAAD, pancreatic adenocarcinoma; PDAC, pancreatic ductal adenocarcinoma; PPI, protein-protein interaction; RFS, recurrence-free survival; TCGA, The Cancer Genome Atlas; UALCAN, University of Alabama Cancer Database.

Electronic supplementary material. The online version of this article (doi: 10.4149/gpb_2021033) contains Supplementary material.

Correspondence to: Dongbo Xue, Department of General Surgery, Key Laboratory of Hepatosplenic Surgery, Ministry of Education, The First Affiliated Hospital of Harbin Medical University, 23 You Zheng Street, Harbin 150001, China
E-mail: xuedongbo@hrbmu.edu.cn

Introduction

Chronic pancreatitis (CP) is a progressive inflammatory disease that causes destruction of pancreatic parenchyma and its replacement by fibrous tissue (Lew et al. 2017; Georg et al. 2020). The continuous inflammatory state of CP promotes the acceleration of tissue repair and may lead to the formation of pancreatic cancer (Takahashi et al. 2020). Pancreatic cancer is one of the most serious malignant tumors in the world, and the 5-year survival rate of pancreatic cancer is still below 5%. Pancreatic ductal adenocarcinoma (PDAC) is the most important and aggressive type of pancreatic cancer, and its clinical manifestations are higFighly similar to those of CP (McGuire 2016; Narkhede et al. 2019). Related research has shown that the risk of pancreatic cancer has a nearly eight-fold increase within 5 years after the diagnosis of CP (Kirkegard et al. 2017). Therefore, elucidating the molecular mechanism governing CP and identifying its relationship with pancreatic cancer may facilitate the treatment of CP and the early prevention of tumor formation and progression.

Gene mutation or abnormal expression is also closely related to the pathogenesis of CP. For example, mutations in CFTR (cystic fibrosis trans-membrane conductance regulator) and SPINK1 (Kazal type 1) are risk factors for lowering the threshold of pancreatitis, and CFTR mutations also increase the risk of pancreatic cancer (Schneider et al. 2011; Jalaly et al. 2017; Cazacu et al. 2018). Studies found that mutations in the CFTR were involved in the process of cancer. First, mutations in the CFTR gene lead to CP, and the long-term inflammation increases the risk of tumor

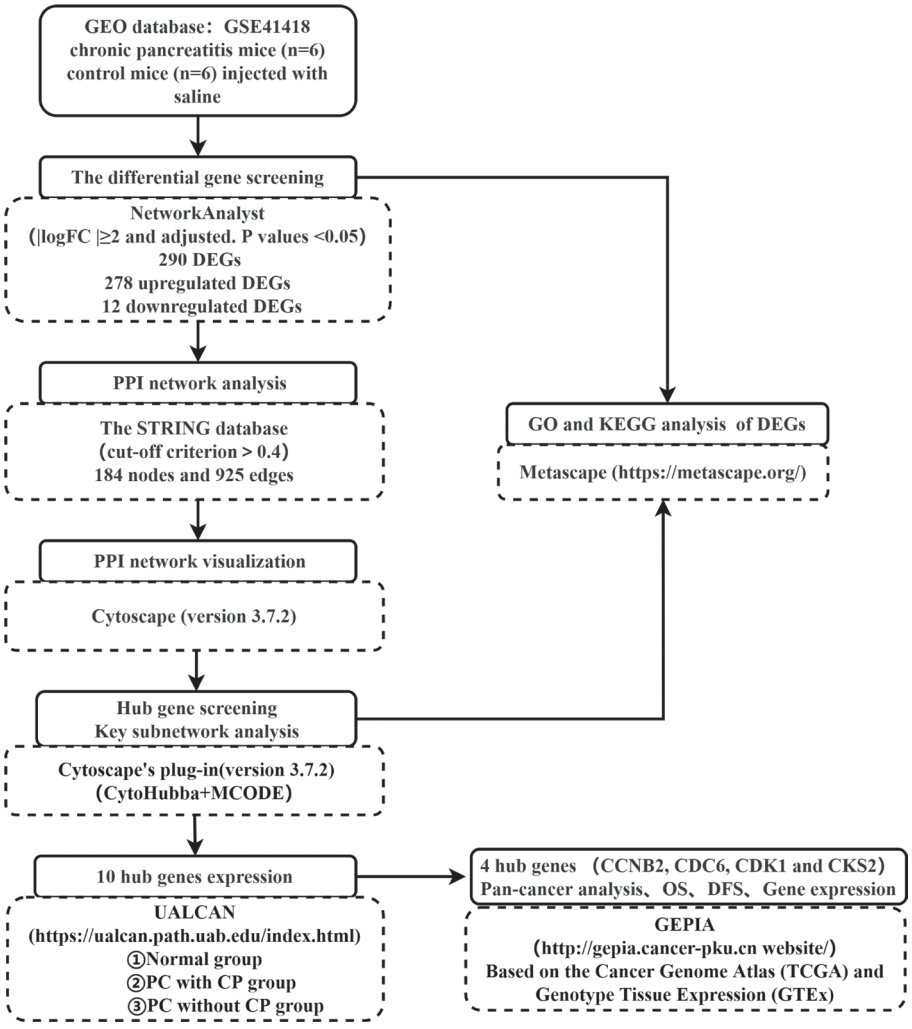


Figure 1. Flow chart of data preparation, processing, analysis, and validation. Dashed lines indicate the basic information of the method or software used. GEO, Gene Expression Omnibus; GO, Gene ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEGs, differentially expressed genes; PPI, protein-protein interaction; GEPIA, Gene Expression Profiling Interactive Analysis; OS, overall survival; RFS, recurrence-free survival; MCODE, Molecular Complex Detection.

transformation (Cazacu et al. 2018). Second, the association of CFTR F508del carriers with a variety of cancers may have potential biological and clinical significance (Shi et al. 2021). Third, CFTR has a low trend of potentially pathogenic variant of uncertain significance in 35% of PDAC cases (Earl et al. 2020). In addition, patients with CFTR mutations and subsequent pancreatic exocrine insufficiency may suffer from deficiencies of selenium and vitamin E. These two substances are antioxidants and are thought to provide cancer protection (Neglia et al. 1995). In addition to gene mutations, it is unknown whether the differentially expressed genes in CP are also involved in the process of cancer. Genes specifically expressed in other diseases also exhibit potential as diagnostic biomarkers and therapeutic targets. In addition, with the rapid development of genetic technology, gene chips and high-throughput sequencing make it possible to study the pathogenesis of various diseases at the molecular level and to use bioinformatic analysis to screen disease-related genes. For example, in cancer research, applying these methods can help us to identify tumor suppressor genes and oncogenes related to tumorigenesis and development and to establish a foundation for further cancer-related research (Xiao et al. 2018; Ni et al. 2019). However, compared with cancer, there are few studies on CP and the association between CP and pancreatic cancer.

In this study, we used microarray data sets in the Gene Expression Omnibus (GEO) database for bioinformatic analysis. By comparing the gene expression in chronic pancreatitis tissue and normal pancreas tissue in two commonly used animal models of CP, we screened out differentially expressed genes (DEGs) and employed the intersection of the two groups of DEGs to select common DEGs. The microarray data was published in 2013 by Neuschwander-Tetri's team (Ulmasov et al. 2013). They found that the genes identified by this analysis were involved mainly in eight cellular pathways: cell cycle, DNA replication and repair, focal adhesion, fibrosis, ubiquitin-mediated proteolysis, cancer, apoptosis, and immune response. The team focused on the description of DEGs related to fibrosis included Mmp7, Pcolce2, Itih4, Wdfy1, and Vtn (Ulmasov et al. 2013). In contrast, we focused on the description of DEGs related to cell cycle, DNA replication and repair, and cancer, and then analyzed the relationship between CP and pancreatic cancer. Then, we performed gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of DEGs, constructed a protein-protein interaction network and discovered hub genes. Later, to explore the connection between CP and pancreatic cancer at the molecular level, we used UALCAN and Gene Expression Profiling Interactive Analysis (GEPIA) online analysis tools for expression and prognosis analysis (see Fig. 1). The results of this study can provide a reference for the diagnosis, treatment, and prognostic evaluation of CP and pancreatic cancer.

Materials and Methods

Gene expression profile data

Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo) (Barrett et al. 2007) is a public genomic database of gene chips and microarrays. The gene expression dataset GSE41418 (Ulmasov et al. 2013) was downloaded from the GEO database. GSE41418 is a mouse CP dataset that includes two different mouse strains commonly used to reproduce CP and human diseases: Harlan mice and Jackson mice. In this dataset, CP was induced in 8-week-old B6J and B6N female mice (18–20 g body weight) by repeated intraperitoneal injections of cerulein (50 mg/kg hourly). Injections were administered every 6 h every other day, and a total of three treatments were administered. Sex- and age-matched control mice received similar saline injections, and to resolve the acute changes, mice were euthanized by CO_2 asphyxiation 3 days after their final cerulein treatment. Pancreatic tissues were used for expression profile chip detection, and saline-injected mice were used as controls. The dataset contained CP Harlan mice (CP-Harlan group, $n = 6$), CP Jackson mice (CP-Jackson group, $n = 6$) and control mice (Control group, $n = 6$) injected with saline.

Identification of DEGs

Network Analyst (http://www.networkanalyst.ca) (Xia et al. 2013; Zhou et al. 2019a), an online tool based on R language, was used to normalize the above datasets and analyze DEGs between normal pancreatic tissues and CP tissues. To identify genes with statistically significant differential expression, the false discovery rate (FDR)-corrected p-value (adjusted p-value) < 0.05, and |log2 fold-change (FC)| ≥ 2 were selected as the screening thresholds. DEGs and their adjusted p-values were calculated by the R"limma"package. Next, the intersection of the two datasets was determined using an online tool for drawing Venn diagrams https://bioinfogp.cnb.csic.es/tools/venny/index.html).

Functional and pathway enrichment analysis of DEGs

To understand the function of these DEGs, Metascape (https://metascape.org/), a free gene annotation and analysis resource, was used to perform GO enrichment and gene enrichment and KEGG pathway enrichment analysis (Zhou et al. 2019b). $p < 0.01$ was considered to be a significant enrichment of DEGs.

PPI network and module analysis

The STRING (https://string-db.org/) (version 11.0, Szklarczyk et al. 2019) is a database of known and predicted protein-

protein interactions (PPI). The interactions include direct (physical) and indirect (functional) associations. Therefore, the database was used to build a PPI network and it was visualized using Cytoscape software (version 3.7.2, Shannon et al. 2003). The number of interactions between proteins was screened using the confidence level. When the number of interactions reached 530'027'879, the interaction with medium or higher confidence was predicted (the comprehensive score ≥ 0.4). Thus, the cut-off criterion was set to a comprehensive score of ≥ 0.4. In addition, CytoHubba of Cytoscape (Chin et al. 2014) was used to calculate the top 10 hub genes using the MCC algorithm. The hub gene plays a vital role in biological processes. An interaction network can be constructed through co-expression or protein interactions, and the key gene can then be screened according to the network topology. Molecular Complex Detection (MCODE) (Bader and Hogue 2003) is a plug-in in Cytoscape for clustering and building functional modules in large protein interaction networks. The software calculated the MCODE score and selected the significant modules of key genes using the screening criteria of Degree Cutoff = 2, Haircut on, Node Score Cut-off = 0.2, k-core = 2, and Max depth = 100. Detailed procedures for the MCODE algorithm are described in reference (Bader and Hogue 2003).

Clinical analyses of hub genes

UALCAN (https://ualcan.path.uab.edu/index.html) (Chandrashekar et al. 2017) is an online portal for facilitating tumor subgroup gene expression based on The Cancer Genome Atlas (TCGA) data. We used the database to analyze the expression of screened hub genes in normal pancreatic tissues, pancreatic adenocarcinoma (PAAD) with CP, and PAAD without CP, and identified four candidate genes. Detailed

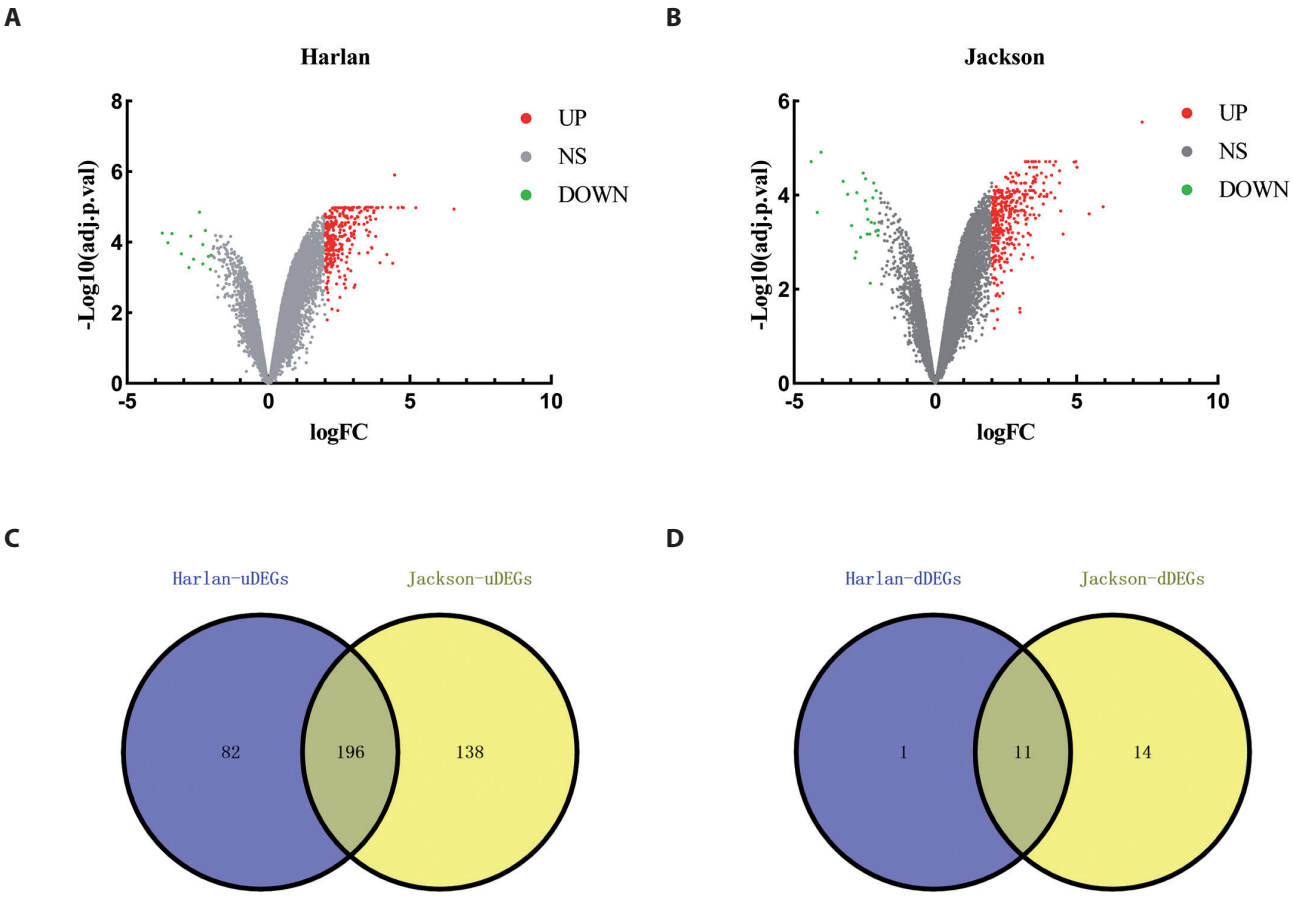


Figure 2. Identification of differentially expressed genes: GSE41418-C57BL/6NHsd (Harlan; **A**) and GSE41418-C57BL/6J (Jackson; **B**). Red color indicates uDEGs (up differentially expressed genes; UP), blue color represents dDEGs (down differentially expressed genes; DOWN), and gray color represents genes that are not significantly different in expression (NS). The criterion: |Foldchange| > 2, adjusted $p < 0.05$ was determined as the cutoff value. Venn diagram shows DEGs shared by GSE41418-C57BL/6NHsd (Harlan) and GSE41418-C57BL/6J (Jackson) uDEGs (**C**) and dDEGs (**D**). A total of 196 uDEGs and 11 dDEGs were identified in the intersections. logFC, log2 fold-change; Adj.p.val, adjusted *p*-value.

procedures for UALCAN are described in reference (Chandrashekar et al. 2017). Here, we provide a brief introduction: enter the gene symbols of the target genes after entering the website and select the corresponding disease dataset; click explore, select the expression option, and select corresponding options according to the research purpose.

GEPIA (http://gepia.cancer-pku.cn website/) (Tang et al. 2017) is a dataset based on TCGA and Genotype Tissue Expression (GTEx), which can provide a fast and customizable network tool. Hence, we utilized the database to analyze the expression of four candidate genes in pan-cancer and selected the default parameters. Detailed procedures for GEPIA are described in reference (Tang et al. 2017). Here, we give a brief introduction: enter the gene symbols of the target genes after entering the website and select "Expression DIY"; after adding all tumor databases, click on the plot to perform pan-cancer expression analysis of the target genes.

A single-factor analysis of variance was conducted to identify significant differences among the groups. Subsequently, the expression levels, overall survival (OS), and recurrence-free survival (RFS) of the four candidate genes in pancreatic cancer were analyzed using GEPIA. The specific operation method is described above.

Results

DEGs between the CP group and control group

When adjusted for $p < 0.05$ and $|\log FC| \geq 2$, a total of 290 DEGs were detected in the CP-Harlan group and the Control group, including 278 upregulated DEGs and 12 downregulated DEGs. A total of 359 DEGs were detected in the CP-Jackson group and the Control group, 334 of which were upregulated and 25 of which were downregulated. Volcano diagrams for the CP-Harlan group and CP-Jackson group *vs.* the Control group are shown in Figure 2A and B, respectively. The intersection of the differential genes obtained from the two groups revealed that 196 DEGs were upregulated and 11 DEGs were downregulated in both CP and control mice (Fig. 2C and D).

GO enrichment analysis and KEGG pathway analysis of DEGs

Metascape for GO enrichment and KEGG pathway enrichment analysis was used to further study the biological functions of each DEG. Analysis of the upregulated DEGs biological process (BP) showed that genes were significantly enriched in response to wounding, extracellular structure organization, cell cycle phase transition, DNA replication initiation, negative regulation of immune system process, and positive regulation of cytokine production (Fig. 3A). Regarding molecular function (MF) analysis, genes were significantly enriched in DNA helicase activity, extracellular matrix structural constituent, and cyclin-dependent protein serine/threonine kinase regulator activity (Fig. 3B). Cell composition (CC) analysis showed that the genes were primarily enriched in collagen-containing extracellular matrix, vacuolar lumen, collagen trimer, cyclin-dependent protein kinase holoenzyme complex, and the MCM complex (Fig. 3C). Due to the small number of downregulated DEGs, no significant enrichment results were obtained. In addition, KEGG pathway analysis showed that upregulation of DEGs was closely related to the cell cycle, DNA replication, phagosome, and *Staphylococcus aureus* infection (Fig. 3D).

PPI network construction and hub gene recognition

A total of 207 DEGs were analyzed using the STRING software. After visualization and analysis using Cytoscape, interactions with a score higher than 0.4, were screened. The PPI network consisted of 184 nodes and 925 edges, including 196 upregulated and 11 downregulated genes (Fig. 4A). CytoHubba was used with the MCC algorithm to calculate the top 10 hub genes, including CDK1, CCNA2, RFC4, MCM7, MCM4, CDC6, MCM5, MAD2L1, CKS2, and CCNB2 (Fig. 4B, Table 1). Subsequently, MCODE was used to cluster and construct functional modules on the large gene (protein) network obtained and dig out protein complexes or corresponding functional modules. Cluster 1, which had the highest score, was obtained from the PPI network using MCODE, including 27 nodes and 319 edges (Fig. 4C, Table 2). Enrichment analysis of the GO and KEGG pathways showed that the genes in this cluster were primarily related to the cell cycle and DNA replication (Fig. 5A and B). In addition, we found that the top 10 hub genes were also in cluster 1 (Fig. 5C and D).

Table 1. Top 10 hub genes ranked by MCC algorithm in PPI network

Rank	Gene
1	CDK1
2	CCNA2
3	MCM7
4	MCM4
5	RFC4
6	MCM5
7	CDC6
8	MAD2L1
9	CKS2
10	CCNB2

PPI, protein-protein interaction. The score is 1.43×10^{17}.

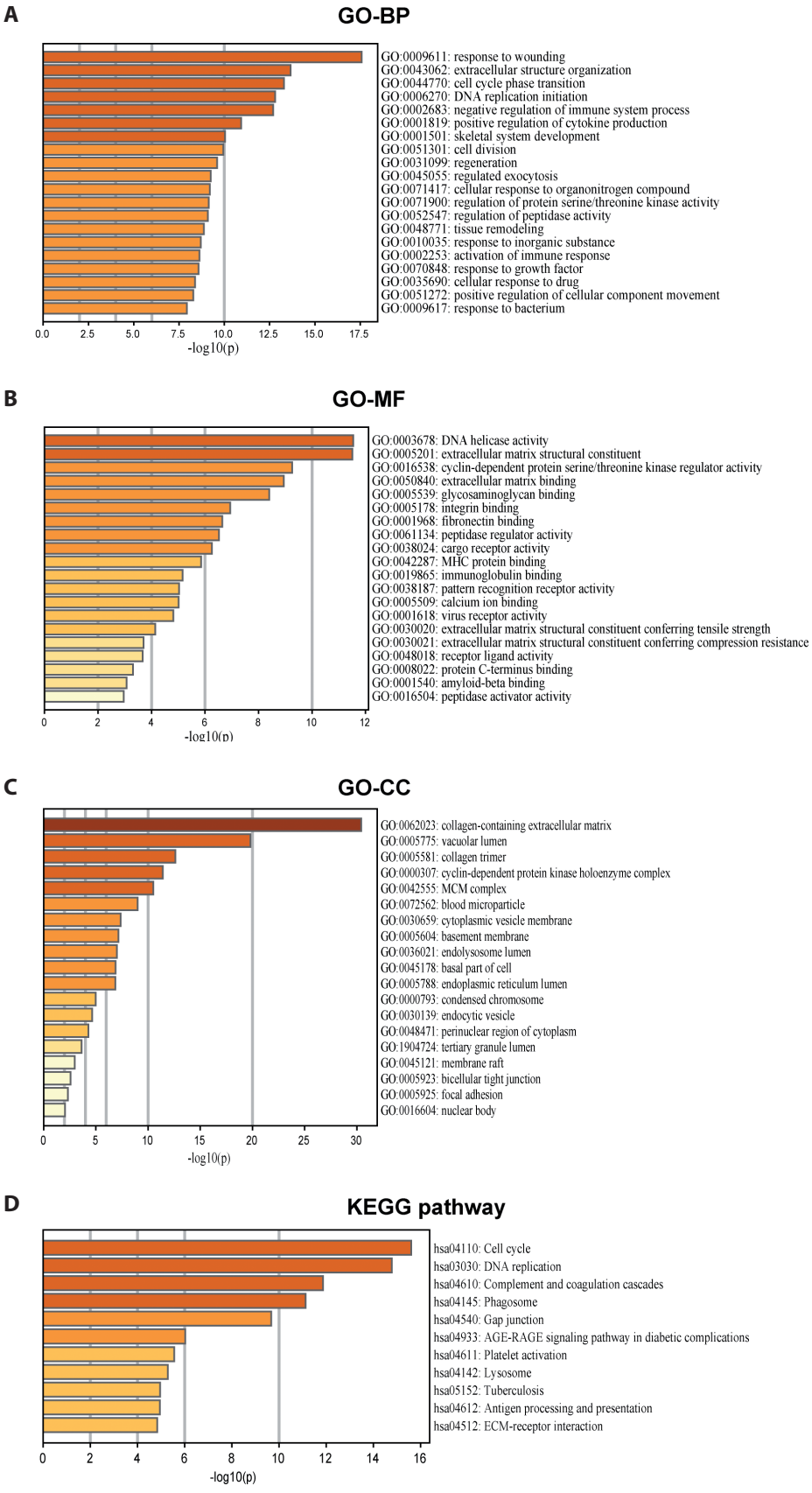


Figure 3. GO and KEGG pathway enrichment analysis of upregulated genes. Enrichment of BP (biological process; **A**), MF (molecular function; **B**), CC (cell composition; **C**) and KEGG (Kyoto Encyclopedia of Genes and Genomes; **D**). GO, Gene ontology.

Table 2. Top 3 clusters from PPI networks analyzed by MCODE

Cluster	Score	Gene
1	24.538	HELLS, CDK4, DUT, SPC25, UBE2C, LIG1, MCM7, ECT2, MCM2, MAD2L1CKS1B, CCNB2, MCM4, CKS2, PRIM1, RFC4, NCAPG2,CCNA2, CDK1, MCM5, CM6, PBK, CDC6, CDT1, MKI67, BIRC5, RAD51
2	11.76	CTGF, FBN1, TYROBP, C1QC, AIF1, MMP2, MPEG1, POSTN, SPP1, LOXL1FCER1G, FCGR3A, C1QB, C1QA, PTPRC, CDH11, LUM, COL3A1, LY86, CD53, COL1A2, COL5A2, CD68, CD74, CTSS, BGN
3	3.778	APOE, GAL, ANXA1, FETUB, AMBP, CD44, GC, LGALS3, ADCY7, ITIH4

PPI, protein-protein interaction; MCODE, molecular complex detection.

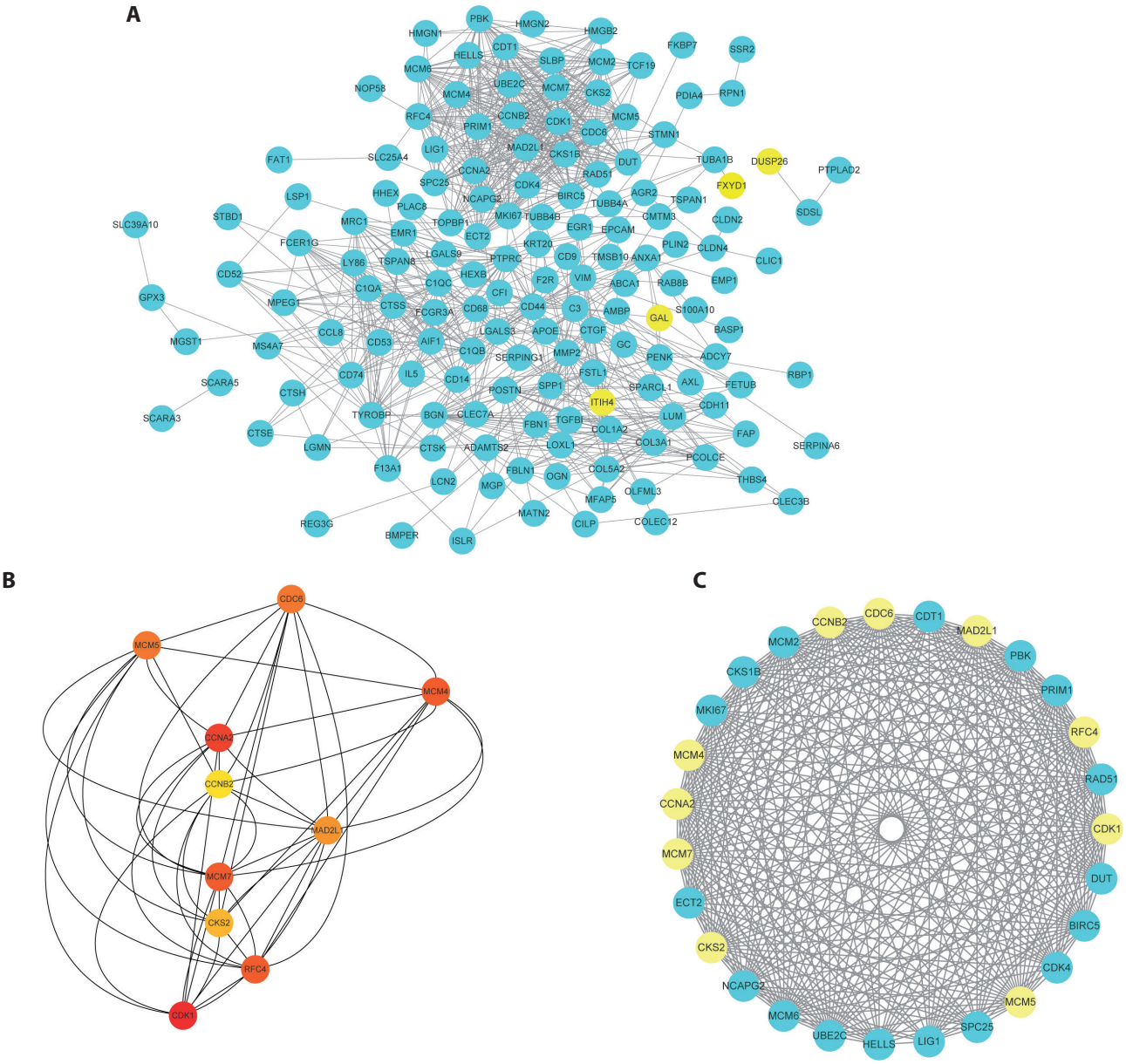


Figure 4. PPI (protein-protein interaction) network and cluster analysis. **A.** PPI network of DEGs generated by STRING (a total of 184 nodes and 925 edges). Blue color indicates uDEGs, yellow color represents dDEGs. **B.** The hub genes calculated with MCC algorithm by cytoHubba. **C.** The most significant cluster generated from the PPI network (yellow color genes are the same as in B). For abbreviations, see Fig. 2.

Association between gene expression and clinical relevance and predictive targets in pancreatic adenocarcinoma with CP

We predicted that many of the screened hub genes acted as oncogenes in tumors; therefore, we further explored the relationship between CP and pancreatic cancer and analyzed the expression of the hub genes in pancreatic cancer. We used UALCAN, which is based on TCGA database, to analyze the expression of CDK1, CCNA2, RFC4, MCM7, MCM4, CDC6, MCM5, MAD2L1, CKS2, and CCNB2 genes in healthy subjects and pancreatic adenocarcinoma patients with and without CP. The results showed that the expression of CCNB2, CDC6, CDK1, and CKS2 was significantly higher in pancreatic adenocarcinoma patients with CP than in healthy subjects (Fig. 6 and Fig. S1 (see Supplementary material). Subsequently, we utilized the GEPIA database to analyze the expression levels of the above four genes in 33 cancer types. The full name of the cancer type abbreviations can be seen in Supplementary Table S1. Analysis of 9,736 tumor samples and 8,587 normal samples from 33 cancers included in TCGA and GTEx projects was conducted. The analysis results showed that CCNB2 was significantly differentially

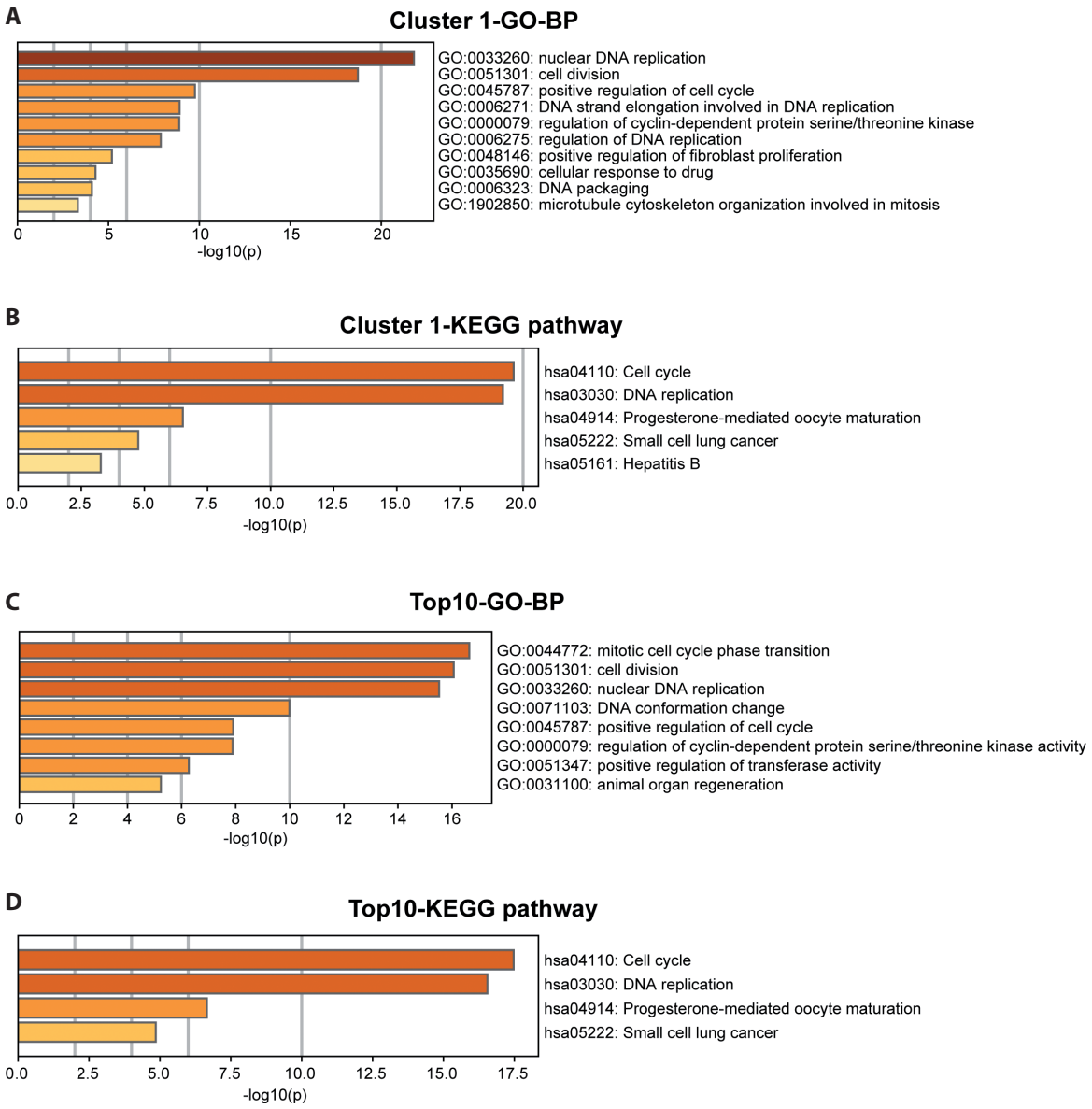


Figure 5. GO and KEGG pathway enrichment analysis of cluster 1 and the top 10 hub genes. Enrichment of BP (**A**), KEGG pathways (**B**) in cluster 1. Enrichment of BP (**C**) and KEGG pathways (**D**) in top 10 hub genes. For abbreviations, see Fig. 3.

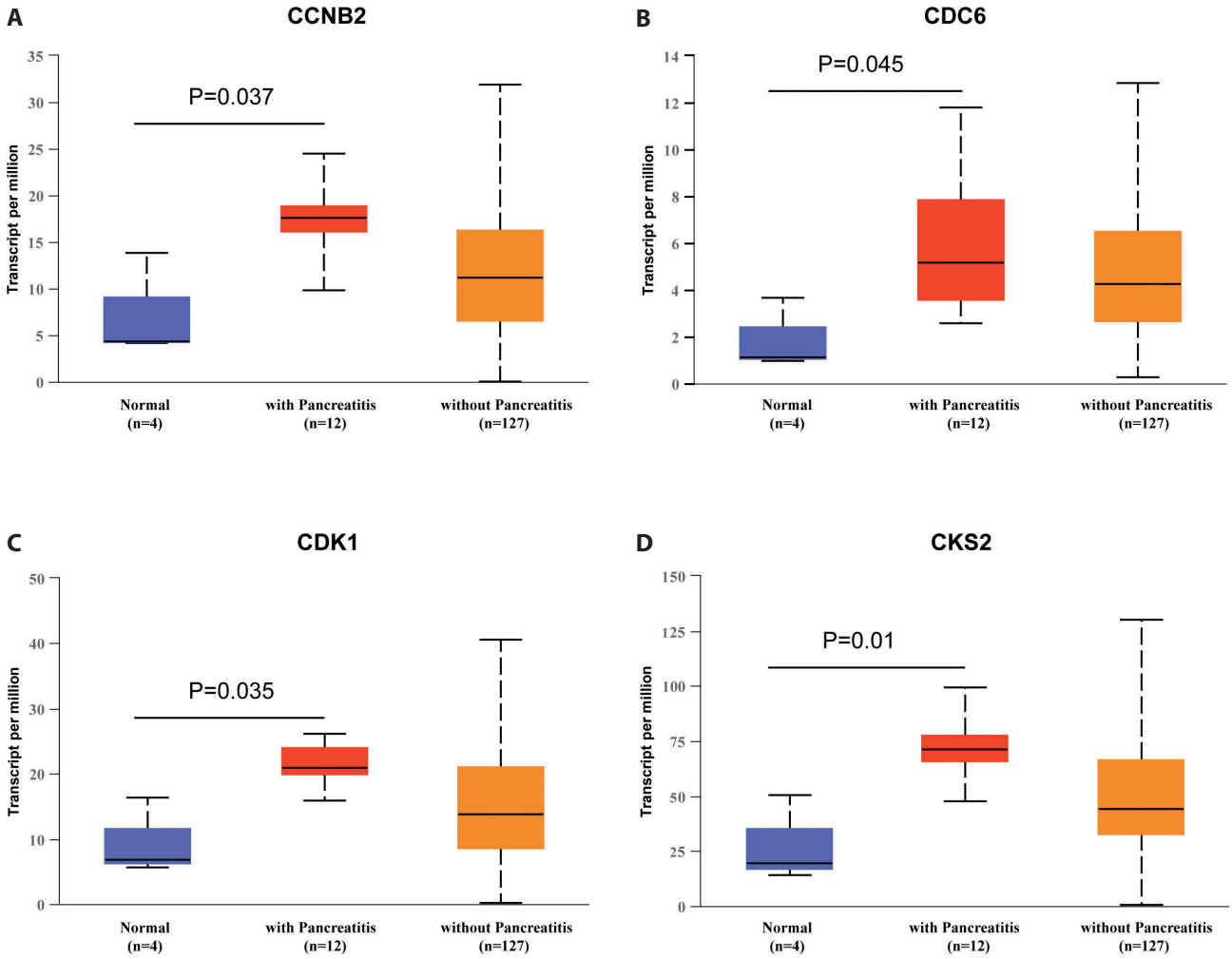


Figure 6. Expression of hub genes CCNB2 (**A**), CDC6 (**B**), CDK1 (**C**), and CKS2 (**D**) in healthy subjects (Normal) and in pancreatic adenocarcinoma patients with and without pancreatitis. The Cancer Genome Atlas (TCGA) database was used.

expressed in 23 cancer types (Fig. 7A and Fig. S2) and CDC6 was significantly differentially expressed in 22 cancer types, including pancreatic cancer (Fig. 7B and Fig. S3). CDK1 was significantly differentially expressed in 22 cancer types, including pancreatic cancer (Fig. 7C and Fig. S4), whereas CKS2 was significantly differentially expressed in 24 cancer types, including pancreatic cancer (Fig. 7D and Fig. S5). The expression of these four genes in the 20 cancer types was significantly different. Among the cancer types, the e expression fold of the genes expression of the genes was only significantly downregulated in acute myeloid leukemia, and the genes were significantly upregulated in other tumors (Fig. 7E). Next, we used the GEPIA online database to analyze the impact of the above four genes on the prognosis of pancreatic cancer and found that high expression levels of CCNB2, CDC6, CDK1, and CKS2 genes were adversely associated with the OS and RFS of pancreatic cancer patients (Figs. 8 and 9). Therefore, CCNB2, CDC6, CDK1, and CKS2 may be potential targets that affect the prognosis of patients with pancreatic cancer accompanied by CP.

Discussion

In this study, we screened DEGs in CP through a series of bioinformatic analysis tools and performed functional and pathway analyses. The results showed that the upregulated genes primarily focused on the cell cycle, DNA replication, and phagosome activity. Because there were too few downregulated genes, no significant results were obtained. Subsequently, we constructed a PPI network and calculated the top 10 hub genes *via* the MCC algorithm, including CDK1, CCNA2, RFC4, MCM7, MCM4, CDC6, MCM5, MAD2L1, CKS2, and CCNB2. Their function and pathway

enrichment are also mainly concentrated in the cell cycle and DNA replication. Increasingly, we found that the functions and pathways involved in the above genes are closely related to tumorigenesis and metastasis. Chuang et al. (2010) reported that relatively small quantitative changes in the expression, steady state, or subcellular distribution of MCM family proteins can affect DNA replication, have different and serious consequences for tumor development, and increase cancer susceptibility. Buccitelli et al. (2017) found that the overexpression of cell cycle-related genes is a feature of proliferation and possible tumor evolution through pan-cancer analysis. Lundberg et al. (2020) have shown that simple measurement of cell cycle activity can yield independent prognostic information on pan-cancer levels.

We validated the expression of these 10 genes in pancreatic cancer based on CP using the UALCAN online tool, and the expression of CCNB2, CDC6, CDK1 and CKS2 in pancreatic adenocarcinoma patients with CP was significantly higher than that in healthy subjects ($p < 0.05$). Next, we used the GEPIA database to analyze the expression levels and prognostic values of the above four genes in pancreatic cancer. The analysis found that compared with normal pancreatic tissue, CCNB2, CDC6, CDK1 and CKS2 genes were significantly overexpressed in pancreatic cancer, and the high expression levels of these four genes were related to poor OS and RFS in patients with pancreatic cancer.

Cyclin B2 (CCNB2) is a member of the cyclin family, especially the type B cyclin. B type cell cycle proteins B1 and

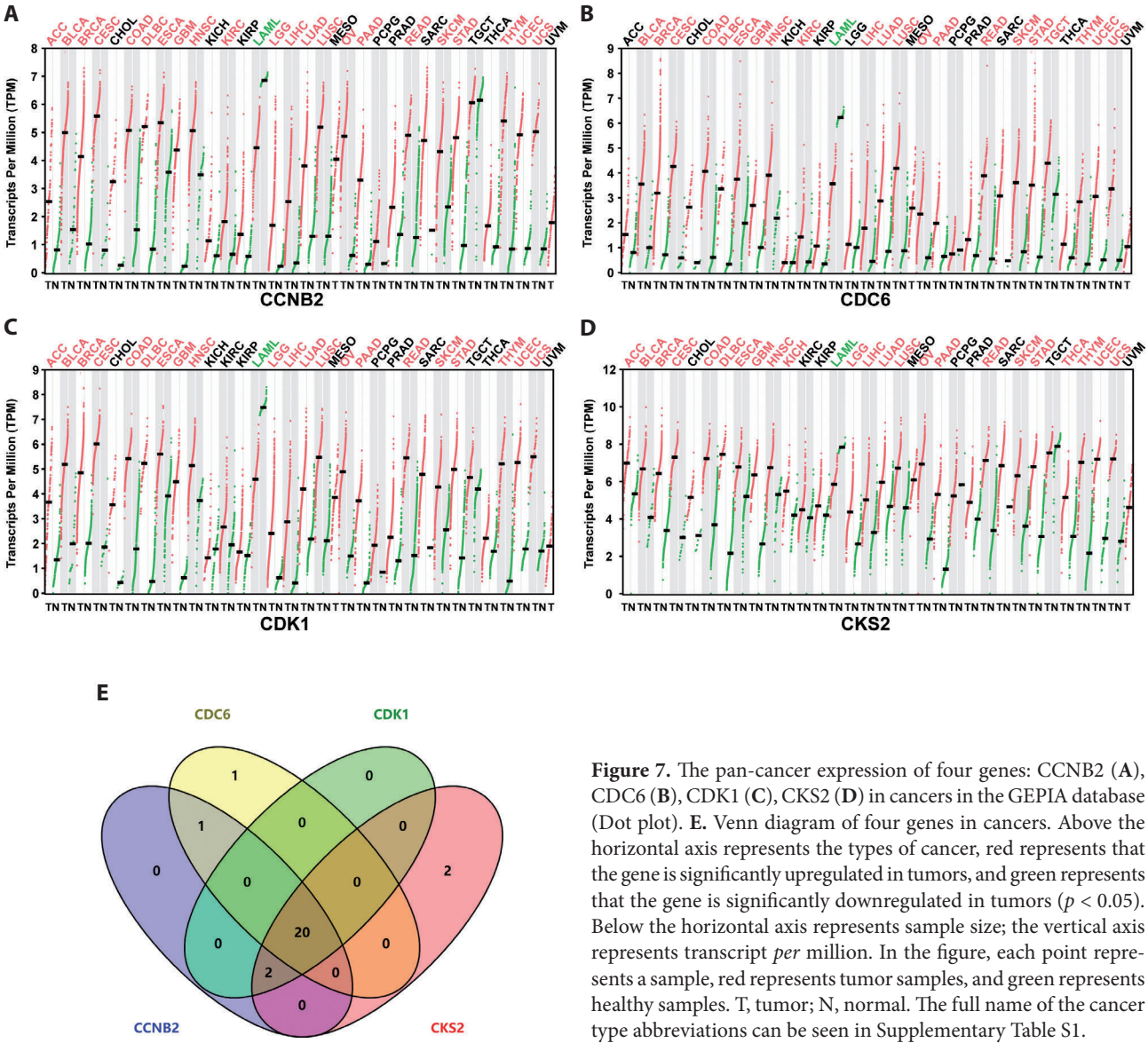


Figure 7. The pan-cancer expression of four genes: CCNB2 (**A**), CDC6 (**B**), CDK1 (**C**), CKS2 (**D**) in cancers in the GEPIA database (Dot plot). **E.** Venn diagram of four genes in cancers. Above the horizontal axis represents the types of cancer, red represents that the gene is significantly upregulated in tumors, and green represents that the gene is significantly downregulated in tumors ($p < 0.05$). Below the horizontal axis represents sample size; the vertical axis represents transcript *per* million. In the figure, each point represents a sample, red represents tumor samples, and green represents healthy samples. T, tumor; N, normal. The full name of the cancer type abbreviations can be seen in Supplementary Table S1.

B2 are related to p34cdc2 and play an important role in cell cycle regulation. Overexpression of cyclin B alters chromosome segregation (Sarafan-Vasseur et al. 2002). Cyclin B2 also binds to transforming growth factor beta RII. Therefore, Cyclin B2 may play a key role in transforming growth factor beta-mediated cell cycle control. Current studies have shown that pancreatic acinar cells have the ability to reenter the cell cycle and proliferate in the event of injury or tissue loss. In this process, acinar cells dedifferentiate and promote acinar-to-ductal metaplasia (Bombardo et al. 2018; Nadella et al. 2020). Therefore, the increased expression of CCNB2 may regulate the cell cycle and proliferation of pancreatic acinar cells, which may promote the progression from pancreatitis to pancreatic cancer. In the study of tumors, the gene plays the role of oncogene in acute myeloid leukemia, hepatocellular carcinoma, bladder cancer and breast cancer (Lei et al. 2016; Li et al. 2019a, 2019b; Naorem et al. 2019). The increased expression of CCNB2 RNA and protein can promote the abovementioned tumor cell proliferation and cell cycle changes. Our results also suggest that the gene may play an important role in the diagnosis and treatment of pancreatic cancer.

Cell division cycle 6 (CDC6), an important regulator of DNA replication, is strongly correlated with pancreatitis (Salabat et al. 2008) and has been expressed in esophageal squamous-cell carcinoma and breast cancer, and high expression levels of CDC6 are associated with poor prognosis in cancer patients (Liu et al. 2010; Ke et al. 2020). Ectopic overexpression of CDC6 leads to apoptosis resistance, cell cycle arrest in G0/G1 phase, epithelial to mesenchymal transformation (EMT), and cell aging caused by DNA overreplication, DNA damage and genome instability (Borlado and Mendez 2008; Yu et al. 2019). To date, studies have shown that inducing apoptosis of pancreatic acinar cells can reduce the degree of acute severe pancreatitis. Due to the resistance of CDC6 to apoptosis, its increased expression level may lead to the proliferation of pancreatic acinar cells and the occurrence of pancreatitis (Bhatia 2004). Therefore, the increase in the CDC6 expression level may play a role in the abovementioned process. Our prediction results show that the CDC6 gene is overexpressed in both chronic pancreatitis and pancreatic cancer tissues. Given the function of CDC6, this has the potential to participate in pancreatic cancer as an oncogene (Lim and Townsend 2020); however, whether its overexpression in chronic pancreatitis plays a role in the progression of the disease to tumor formation and its mechanism have not been elucidated. The molecular mechanism by which CDC6 induces aging also warrants further research.

The protein encoded by the cyclin-dependent kinase 1 (CDK1) gene is a member of the Ser/Thr protein kinase family and is essential for the G1/S and G2/M phase transitions in the eukaryotic cell cycle. To date, it has been reported that CDK1 is overexpressed in various cancers, such as cholangiocarcinoma, colorectal cancer, and renal cell carcinoma, and it affects the prognosis of patients (Yang et al. 2018; Yamamura et al. 2020; Zhu et al. 2020). The expression of CDK1 in pancreatic cancer has been shown to be elevated and to predict poor prognosis, while the mechanisms governing the role of CDK1 in pancreatic cancer have not been elucidated (Dong et al. 2019; Piao et al. 2019). Our research found that CDK1 is enriched in the cell cycle pathway and participates in mitotic cell cycle processes, mitotic cell cycle phase transitions, cell cycle processes, and DNA replication. These processes play important roles in the abnormal proliferation of cancer cells and the regulation of genetic stability. Therefore, our results indicate that CDK1 may play an important role in the development of pancreatic cancer. As also mentioned above, due

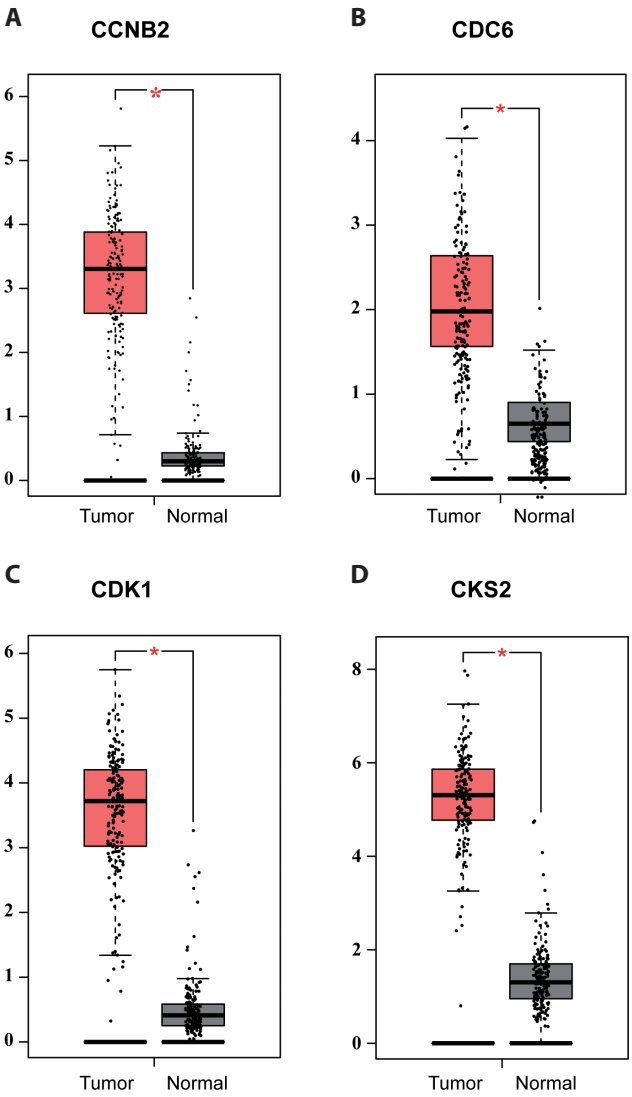


Figure 8. Expression of four hub genes: CCNB2 (**A**), CDC6 (**B**), CDK1 (**C**), CKS2 (**D**) in healthy subject (Normal; the number of samples 171) and in pancreatic cancer patients (Tumor; the number of samples 179). * statistically significant differences, $p < 0.05$.

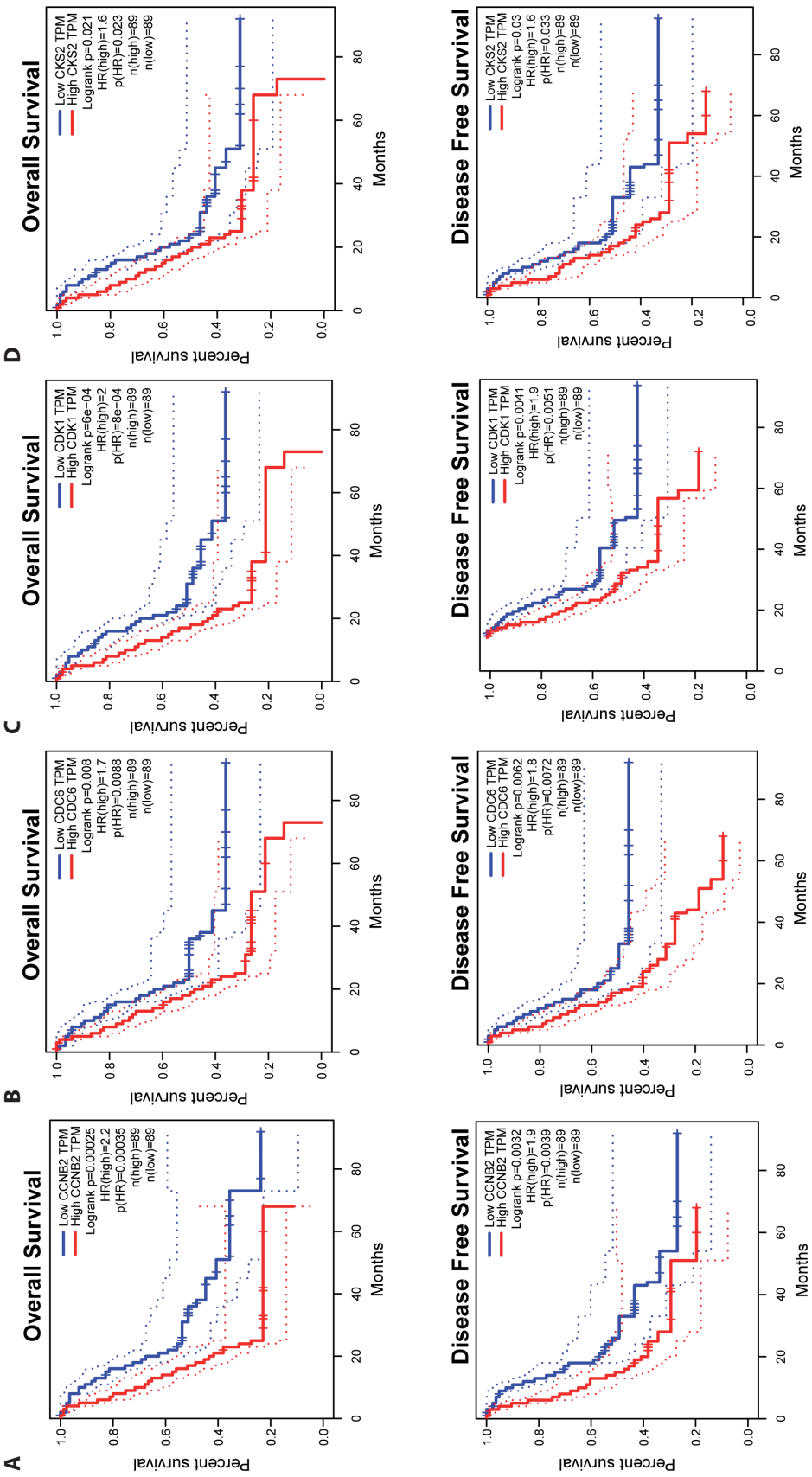


Figure 9. Predictive targets for pancreatic cancer prognosis. The prognostic values of the CCNB2 (**A**), CDC6 (**B**), CDK1 (**C**) and CKS2 (**D**) genes were explored by GEPIA. With high-level expression genes, CCNB2, CDC6, CDK1 and CKS2 patients had a worse overall survival (upper) and disease free survival (bottom) compared with the corresponding low expression level group in pancreatic cancer (logrank test, $p < 0.05$). The hazard ratio (HR) was calculated based on the Cox PH model, and the 95% CI was added as a dotted line. CI, confidence intervals.

Table 3. Possible mechanisms involved in hub genes

Gene	Function	Role in cancers	Related cancers	Possible mechanisms in CP/PC	References
CCNB2	Cell proliferation	Oncogene	LAML LIHC BRCA BLCA	Acinar cells dedifferentiate and promote acinar-to-ductal metaplasia	Bombardo et al. 2018, Nadella et al. 2020
CDC6	EMT Cell aging	Oncogene	ESCA BRCA PAAD	Apoptosis resistance	Salabat et al. 2008, Bhatia 2004, Lim and Townsend 2020
CDK1	Cell proliferation	Oncogene	CHOL COAD KIRC PAAD	Pancreatic fibrosis and acinar-to-ductal metaplasia	Dong et al. 2019, Piao et al. 2019
CKS2	Cell proliferation	Oncogene	LIHC BRCA ESCA	Proliferation of acinar cells and promote acinar-to-ductal metaplasia	Bombardo et al. 2018, Nadella et al. 2020

LAML, acute myeloid leukemia; LIHC, liver hepatocellular carcinoma; BRCA, breast invasive carcinoma; BLCA, cladder urothelial carcinoma; ESCA, esophageal carcinoma; PAAD, pancreatic adenocarcinoma; CHOL, cholangiocarcinoma; COAD, colon adenocarcinoma; KIRC, kidney renal cell carcinoma; EMT, epithelial to mesenchymal transformation; CP, chronic pancreatitis; PC, pancreatic cancer.

to the important role played by CDK1 in the cell cycle, we speculate that its increased expression in CP may accelerate pancreatic fibrosis and acinar-to-ductal metaplasia which, in turn, leads to the development of pancreatic cancer.

CDC28 protein kinase regulatory subunit 2 (CKS2) is a member of the CKS protein family, which is a cyclin-dependent kinase (CDK) binding protein and an important component of cell cycle regulation (Zhang et al. 2019). Currently, increased expression of CKS2 has been found in hepatocellular carcinoma, breast cancer, esophageal squamous-cell carcinoma, and other tumors to promote tumor cell proliferation and invasion and predicts a worse prognosis (Kita et al. 2014; Li et al. 2018; Huang et al. 2019). However, the role of CSK2 in pancreatic cancer has not been found because of its role as an oncogene in other tumors, and the results of our analysis show that the expression of CSK2 in pancreatic cancer is also increased compared with the low expression of CKS2, which also has a worse prognosis. Therefore, CKS2 may be an effective diagnosis and treatment target for pancreatic cancer. Meanwhile, the high expression of CKS2 in CP may affect the cell cycle and proliferation of acinar cells, thereby inducing acinar-to-ductal metaplasia of cells and increasing the risk of pancreatic cancer. Possible mechanisms involved in hub genes can be seen in Table 3.

Transdifferentiation of exocrine pancreatic acinar cells into ductal cells is a common pathogenesis of pancreatitis and pancreatic cancer (Hakobyan et al. 2020). Johnson and colleagues found that the expression of oncogenes in acinar cells can induce acinar-ductal metaplasia and pancreatic intraepithelial neoplasia. After a certain incubation period, acinar cells expressing oncogenes begin to form highly differentiated and fibrotic tumors that metastasize to the lungs and liver (Johnson et al. 2019). Therefore, the above results suggest that these four genes may play important roles in the development of chronic pancreatitis and pancreatic cancer.

Conclusions

In our study, the DEGs in chronic pancreatitis tissue were used as the entry point. After identifying the hub genes, it was found that such genes as CCNB2, CDC6, CDK1 and CKS2 may be the key genes in the development of CP and pancreatic cancer. TCGA and GTEX databases were used to verify the expression of these genes and analyze their effects on prognosis to provide a theoretical basis for the diagnosis, prevention and treatment of CP and pancreatic cancer at the molecular level.

Availability of data and materials. The GSE41418 dataset is available in the Gene Expression Omnibus (GEO) database at https://www.ncbi.nlm.nih.gov/geo/, reference number [PMID： 23845568].

Funding. This study was supported by the National Natural Science Foundation of China (grant no. 81770634).

Conflict of interest. The authors declare that they have no conflicts of interest.

References

Bader GD, Hogue CWV (2003): An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics **4,** 2
https://doi.org/10.1186/1471-2105-4-2

Barrett T, Troup DB, Wilhite SE, Ledoux P, Rudnev D, Evangelista C, Kim IF, Soboleva A, Tomashevsky M, Edgar R (2007): NCBI GEO: mining tens of millions of expression profiles--database and tools update. Nucleic Acids Res. **35,** D760-765 https://doi.org/10.1093/nar/gkl887

Bhatia M (2004): Apoptosis versus necrosis in acute pancreatitis. Am. J. Physiol. Gastrointest. Liver Physiol. **286,** G189-196 https://doi.org/10.1152/ajpgi.00304.2003

Bombardo M, Malagola E, Chen R, Carta A, Seleznik GM, Hills AP, Graf R, Sonda S (2018): Enhanced proliferation of pancreatic acinar cells in MRL/MpJ mice is driven by severe acinar injury but independent of inflammation. Sci. Rep. **8,** 9391 https://doi.org/10.1038/s41598-018-27422-0

Borlado LR, Mendez J (2008): CDC6: from DNA replication to cell cycle checkpoints and oncogenesis. Carcinogenesis **29,** 237-243 https://doi.org/10.1093/carcin/bgm268

Buccitelli C, Salgueiro L, Rowald K, Sotillo R, Mardin BR, Korbel JO (2017): Pan-cancer analysis distinguishes transcriptional changes of aneuploidy from proliferation. Genome Res. **27,** 501-511 https://doi.org/10.1101/gr.212225.116

Cazacu IM, Farkas N, Garami A, Balaskó M, Mosdósi B, Alizadeh H, Gyöngyi Z, Rakonczay Z Jr., Vigh E, Habon T (2018): Pancreatitis-associated genes and pancreatic cancer risk: A systematic review and meta-analysis. Pancreas **47,** 1078-1086 https://doi.org/10.1097/MPA.0000000000001145

Chandrashekar DS, Bashel B, Balasubramanya SAH, Creighton CJ, Ponce-Rodriguez I, Chakravarthi Balabhadrapatruni VSK, Varambally S (2017): UALCAN: A portal for facilitating tumor subgroup gene expression and survival analyses. Neoplasia **19,** 649-658 https://doi.org/10.1016/j.neo.2017.05.002

Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY (2014): cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC Syst. Biol. **8** (Suppl. 4), S11 https://doi.org/10.1186/1752-0509-8-S4-S11

Chuang CH, Wallace MD, Abratte C, Southard T, Schimenti JC (2010): Incremental genetic perturbations to MCM2-7 expression and subcellular distribution reveal exquisite sensitivity of mice to DNA replication stress. PLoS Genet. **6,** e1001110 https://doi.org/10.1371/journal.pgen.1001110

Dong S, Huang F, Zhang H, Chen Q (2019): Overexpression of BUB1B, CCNA2, CDC20, and CDK1 in tumor tissues predicts poor survival in pancreatic ductal adenocarcinoma. Biosci. Rep. **39,** BSR20182306 https://doi.org/10.1042/BSR20182306

Earl J, Galindo-Pumarino C, Encinas J, Barreto E, Castillo ME, Pachón V, Ferreiro R, Rodríguez-Garrote M, González-Martínez S, Ramon Y, Cajal T (2020): A comprehensive analysis of candidate genes in familial pancreatic cancer families reveals a high frequency of potentially pathogenic germline variants. EBioMedicine **53,** 102675 https://doi.org/10.1016/j.ebiom.2020.102675

Georg B, Aida H, Jens W, M LM, Julia M (2020): Chronic pancreatitis. Lancet **396,** 499-512 https://doi.org/10.1016/S0140-6736(20)31318-0

Hakobyan D, Medina C, Dusserre N, Stachowicz ML, Handschin C, Fricain JC, Guillermet-Guibert J, Oliveiraet H (2020): Laser-assisted 3D bioprinting of exocrine pancreas spheroid models for cancer initiation study. Biofabrication **12,** 035001 https://doi.org/10.1088/1758-5090/ab7cb8

Huang N, Wu Z, Hong H, Wang X, Yang F, Li H (2019): Overexpression of CKS2 is associated with a poor prognosis and promotes cell proliferation and invasion in breast cancer. Mol. Med. Rep. **19,** 4761-4769 https://doi.org/10.3892/mmr.2019.10134

Jalaly NY, Moran RA, Fargahi F, Khashab MA, Kamal A, Lennon AM, Walsh C, Makary MA, Whitcomb DC, Yadav D (2017): An evaluation of factors associated with pathogenic PRSS1, SPINK1, CTFR, and/or CTRC genetic variants in patients with idiopathic pancreatitis. Am. J. Gastroenterol. **112,** 1320-1329 https://doi.org/10.1038/ajg.2017.106

Johnson BL, d'Alincourt Salazar M, Mackenzie-Dyck S, D'Apuzzo M, Shih HP, Manuel ER, Diamond DJ (2019): Desmoplasia and oncogene driven acinar-to-ductal metaplasia are concurrent events during acinar cell-derived pancreatic cancer initiation in young adult mice. PLoS One **14,** e0221810 https://doi.org/10.1371/journal.pone.0221810

Ke Y, Guo W, Huang S, Li Y, Guo Y, Liu X, Jin Y, Ma H (2020): RYBP inhibits esophageal squamous cell carcinoma proliferation through downregulating CDC6 and CDC45 in G1-S phase transition process. Life Sci. **250,** 117578 https://doi.org/10.1016/j.lfs.2020.117578

Kirkegard J, Mortensen FV, Cronin-Fenton D (2017): Chronic pancreatitis and pancreatic cancer risk: A systematic review and meta-analysis. Am. J. Gastroenterol. **112,** 1366-1372 https://doi.org/10.1038/ajg.2017.218

Kita Y, Nishizono Y, Okumura H, Uchikado Y, Sasaki K, Matsumoto M, Setoyama T, Tanoue K, Omoto I, Mori S (2014): Clinical and biological impact of cyclin-dependent kinase subunit 2 in esophageal squamous cell carcinoma. Oncol. Rep. **31,** 1986-1992 https://doi.org/10.3892/or.2014.3062

Lei CY, Wang W, Zhu YT, Fang WY, Tan WL (2016): The decrease of cyclin B2 expression inhibits invasion and metastasis of bladder cancer. Urol. Oncol. **34,** 237 e1-10 https://doi.org/10.1016/j.urolonc.2015.11.011

Lew D, Afghani E, Pandol S (2017): Chronic pancreatitis: Current status and challenges for prevention and treatment. Dig. Dis. Sci. **62,** 1702-1712 https://doi.org/10.1007/s10620-017-4602-2

Li H, Tian X, Wang P, Huang M, Xu R, Nie T (2019a): MicroRNA-582-3p negatively regulates cell proliferation and cell cycle progression in acute myeloid leukemia by targeting cyclin B2. Cell. Mol. Biol. Lett. **24,** 66 https://doi.org/10.1186/s11658-019-0184-7

Li R, Jiang X, Zhang Y, Wang S, Chen X, Yu X, Ma J, Huang X (2019b): Cyclin B2 overexpression in human hepatocellular carcinoma is associated with poor prognosis. Arch. Med. Res. **50,** 10-17 https://doi.org/10.1016/j.arcmed.2019.03.003

Li Z, Xue TQ, Yang C, Wang YL, Zhu XL, Ni CF (2018): EGFL7 promotes hepatocellular carcinoma cell proliferation and inhibits cell apoptosis through increasing CKS2 expression by activating Wnt/beta-catenin signaling. J. Cell. Biochem. **119,** 10327-10337 https://doi.org/10.1002/jcb.27375

Lim N, Townsend PA (2020): Cdc6 as a novel target in cancer: Oncogenic potential, senescence and subcellular localization. Int. J. Cancer **147,** 1528-1534 https://doi.org/10.1002/ijc.32900

Liu Y, Hock JM, Sullivan C, Fang G, Cox AJ, Davis KT, Davis BH, Li X (2010): Activation of the p38 MAPK/Akt/ERK1/2 signal pathways is required for the protein stabilization of CDC6 and cyclin D1 in low-dose arsenite-induced cell proliferation. J. Cell. Biochem. **111,** 1546-1555 https://doi.org/10.1002/jcb.22886

Lundberg A, Lindstrom LS, Parker JS, Löverli E, Perou CM, Bergh J, Tobin NP (2020): A pan-cancer analysis of the frequency of DNA alterations across cell cycle activity levels. Oncogene **39,** 5430-5440 https://doi.org/10.1038/s41388-020-1367-4

McGuire S (2016): World Cancer Report 2014. Geneva, Switzerland: World Health Organization, International Agency for Research on Cancer, WHO Press, 2015. Adv. Nutr. **7,** 418-419 https://doi.org/10.3945/an.116.012211

Nadella S, Ciofoaia V, Cao H, Kallakury B, Tucker RD, Smith JP (2020): Cholecystokinin receptor antagonist therapy decreases inflammation and fibrosis in chronic pancreatitis. Dig. Dis. Sci. **65,** 1376-1384 https://doi.org/10.1007/s10620-019-05863-5

Naorem LD, Muthaiyan M, Venkatesan A (2019): Integrated network analysis and machine learning approach for the identification of key genes of triple-negative breast cancer. J. Cell. Biochem. **120,** 6154-6167 https://doi.org/10.1002/jcb.27903

Narkhede RA, Desai GS, Prasad PP, Wagle PK (2019): diagnosis and management of pancreatic adenocarcinoma in the background of chronic pancreatitis: Core issues. Dig. Dis. **37,** 315-324 https://doi.org/10.1159/000496507

Neglia JP, FitzSimmons SC, Maisonneuve P, Schöni MH, Schöni-Affolter F, Corey M, Lowenfelset AB (1995): The risk of cancer among patients with cystic fibrosis. Cystic Fibrosis and Cancer Study Group. N. Engl. J. Med. **332,** 494-499 https://doi.org/10.1056/NEJM199502233320803

Ni W, Zhang S, Jiang B, Ni R, Xiao M, Lu C, Liu J, Qu L, Ni H, Zhang W, Zhou P (2019): Identification of cancer-related gene network in hepatocellular carcinoma by combined bioinformatic approach and experimental validation. Pathol. Res. Pract. **215,** 152428 https://doi.org/10.1016/j.prp.2019.04.020

Piao J, Zhu L, Sun J, Li N, Dong B, Yang Y, Chen L (2019): High expression of CDK1 and BUB1 predicts poor prognosis of pancreatic ductal adenocarcinoma. Gene **701,** 15-22 https://doi.org/10.1016/j.gene.2019.02.081

Salabat MR, Melstrom LG, Strouch MJ, Ding XZ, Milam BM, Ujiki MB, Chen C, Pelling JC, Rao S, Grippo PJ (2008): Geminin is overexpressed in human pancreatic cancer and downregulated by the bioflavanoid apigenin in pancreatic cancer cell lines. Mol. Carcinog. **47,** 835-844 https://doi.org/10.1002/mc.20441

Sarafan-Vasseur. N, Lamy A, Bourguignon J, Le Pessot F, Hieter P, Sesboüé R, Bastard C, Frébourg T, Flaman JM (2002): Overexpression of B-type cyclins alters chromosomal segregation. Oncogene **21,** 2051-2057 https://doi.org/10.1038/sj.onc.1205257

Schneider A, Larusch J, Sun X, Aloe A, Lamb J, Hawes R, Cotton P, Brand RE, Anderson MA, Money ME, et al. (2011): Combined bicarbonate conductance-impairing variants in CFTR and SPINK1 variants are associated with chronic pancreatitis in patients without cystic fibrosis. Gastroenterology **140,** 162-171 https://doi.org/10.1053/j.gastro.2010.10.045

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T (2003): Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. **13,** 2498-2504 https://doi.org/10.1101/gr.1239303

Shi Z, Wei J, Na R, Kyle Resurreccion W, Zheng SL, Hulick PJ, Helfand BT, Talamonti MS, Xu J (2021): Cystic fibrosis F508del carriers and cancer risk: Results from the UK Biobank. Int. J. Cancer **148,** 1658-1664 https://doi.org/10.1002/ijc.33431

Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, et al. (2019): STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. **47,** D607-D613 https://doi.org/10.1093/nar/gky1131

Takahashi R, Macchini M, Sunagawa M, Jiang Z, Tanaka T, Valenti G, Renz BW, White RA, Hayakawa Y, Westphalen CB, et al. (2020): Interleukin-1beta-induced pancreatitis promotes pancreatic ductal adenocarcinoma via B lymphocyte-mediated immune suppression. Gut **70,** 330-341 https://doi.org/10.1136/gutjnl-2019-319912

Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z (2017): GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. **45,** W98-102 https://doi.org/10.1093/nar/gkx247

Ulmasov B, Oshima K, Rodriguez MG, Cox RD, Neuschwander-Tetri BA (2013): Differences in the degree of cerulein-induced chronic pancreatitis in C57BL/6 mouse substrains lead to new insights in identification of potential risk factors in the development of chronic pancreatitis. Am. J. Pathol. **183,** 692-708 https://doi.org/10.1016/j.ajpath.2013.05.020

Xia J, Lyle NH, Mayer ML, Pena OM, Hancock RE (2013): INVEX--a web-based tool for integrative visualization of expression data. Bioinformatics **29,** 3232-3234 https://doi.org/10.1093/bioinformatics/btt562

Xiao H, Xu D, Chen P, Zeng G, Wang X, Zhang X (2018): identification of five genes as a potential biomarker for predicting progress and prognosis in adrenocortical carcinoma. J. Cancer **9,** 4484-4495 https://doi.org/10.7150/jca.26698

Yamamura M, Sato Y, Takahashi K, Sasaki M, Harada K (2020): The cyclindependent kinase pathway involving CDK1 is a potential therapeutic target for cholangiocarcinoma. Oncol. Rep. **43,** 306-317 https://doi.org/10.3892/or.2019.7405

Yang CA, Huang HY, Yen JC, Chang JG (2018): Prognostic value of RNASEH2A-, CDK1-, and CD151-related pathway gene profiling for kidney cancers. Int. J. Mol. Sci. **19,** 1586 https://doi.org/10.3390/ijms19061586

Yu X, Liu Y, Yin L, Peng Y, Peng Y, Gao Y, Yuan B, Zhu Q, Cao T, Xie B, et al. (2019): Radiation-promoted CDC6 protein stability contributes to radioresistance by regulating senescence and epithelial to mesenchymal transition. Oncogene **38,** 549-563 https://doi.org/10.1038/s41388-018-0460-4

Zhang J, Song Q, Liu J, Lu L, Xu Y, Zheng W (2019): Cyclin-dependent kinase regulatory subunit 2 indicated poor prognosis and facilitated aggressive phenotype of hepatocellular carcinoma. Dis. Markers **2019,** 8964015 https://doi.org/10.1155/2019/8964015

Zhou G, Soufan O, Ewald J, Hancock REW, Basu N, Xia J (2019a): NetworkAnalyst 3.0: a visual analytics platform for comprehensive gene expression profiling and meta-analysis. Nucleic Acids Res. **47,** W234-W241 https://doi.org/10.1093/nar/gkz240

Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK (2019b): Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. **10,** 1523 https://doi.org/10.1038/s41467-019-09234-6

Zhu Y, Bian Y, Zhang Q, Hu J, Li L, Yang M, Qian H, Yu L, Liu B, Qian X (2020): LINC00365 promotes colorectal cancer cell progression through the Wnt/beta-catenin signaling pathway. J. Cell. Biochem. **121,** 1260-1272 https://doi.org/10.1002/jcb.29359

Received: May 6, 2021
Final version accepted: August 18, 2021

doi: 0.4149/gpb_2021033

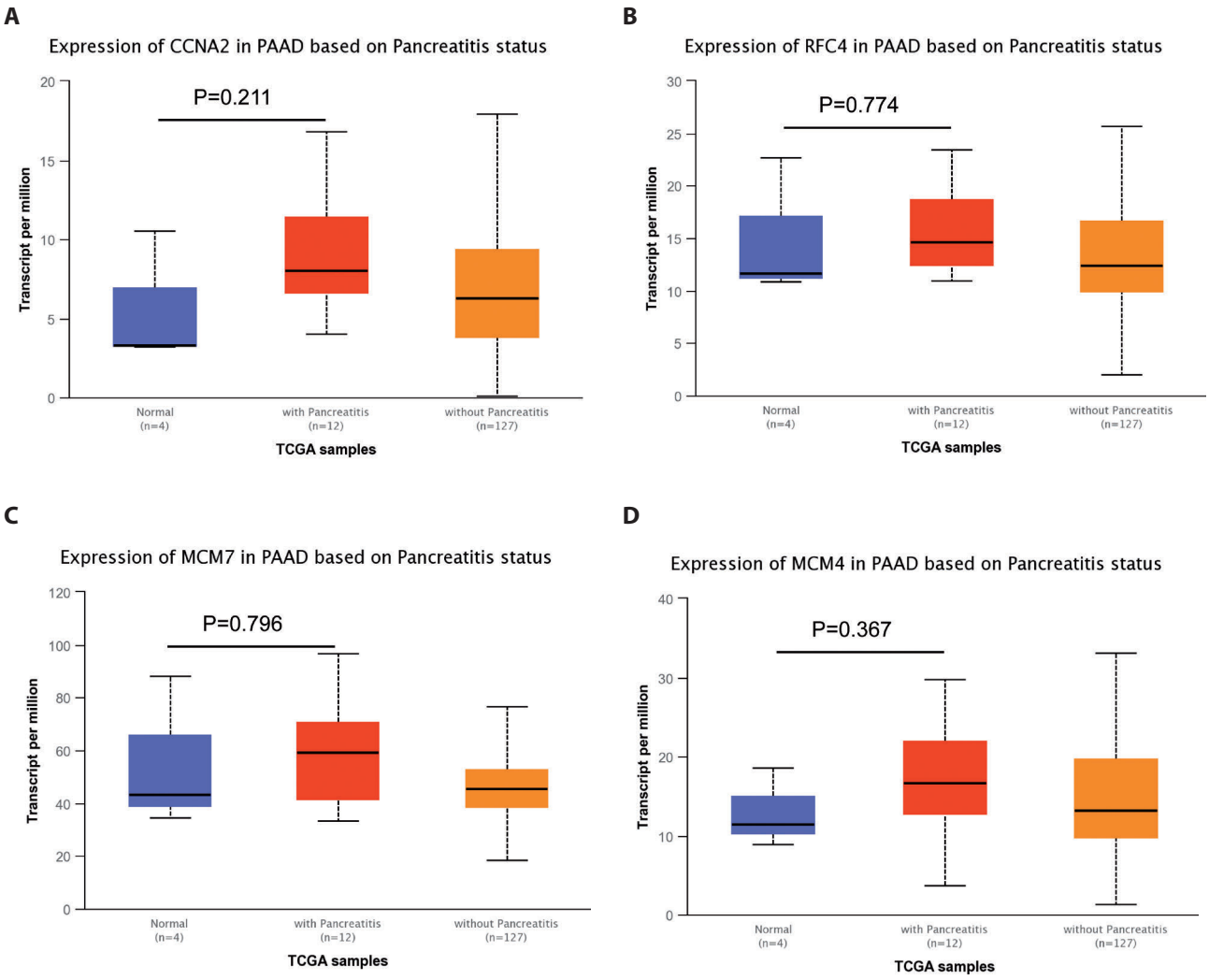
Supplementary Material

Identification of hub genes in chronic pancreatitis and analysis of association with pancreatic cancer *via* bioinformatic analysis

Hongxuan Li, Chenjun Hao, Qiu Yang, Wenqi Gao, Biao Ma and Dongbo Xue

Department of General Surgery, Key Laboratory of Hepatosplenic Surgery, Ministry of Education, The First Affiliated Hospital of Harbin Medical University, Harbin, China

Supplementary Figures



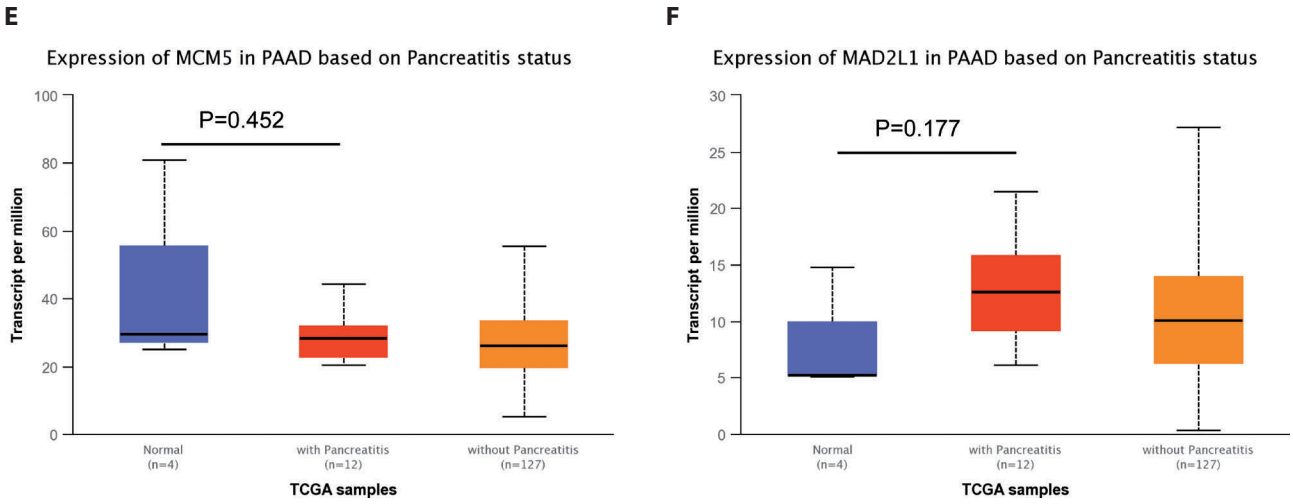
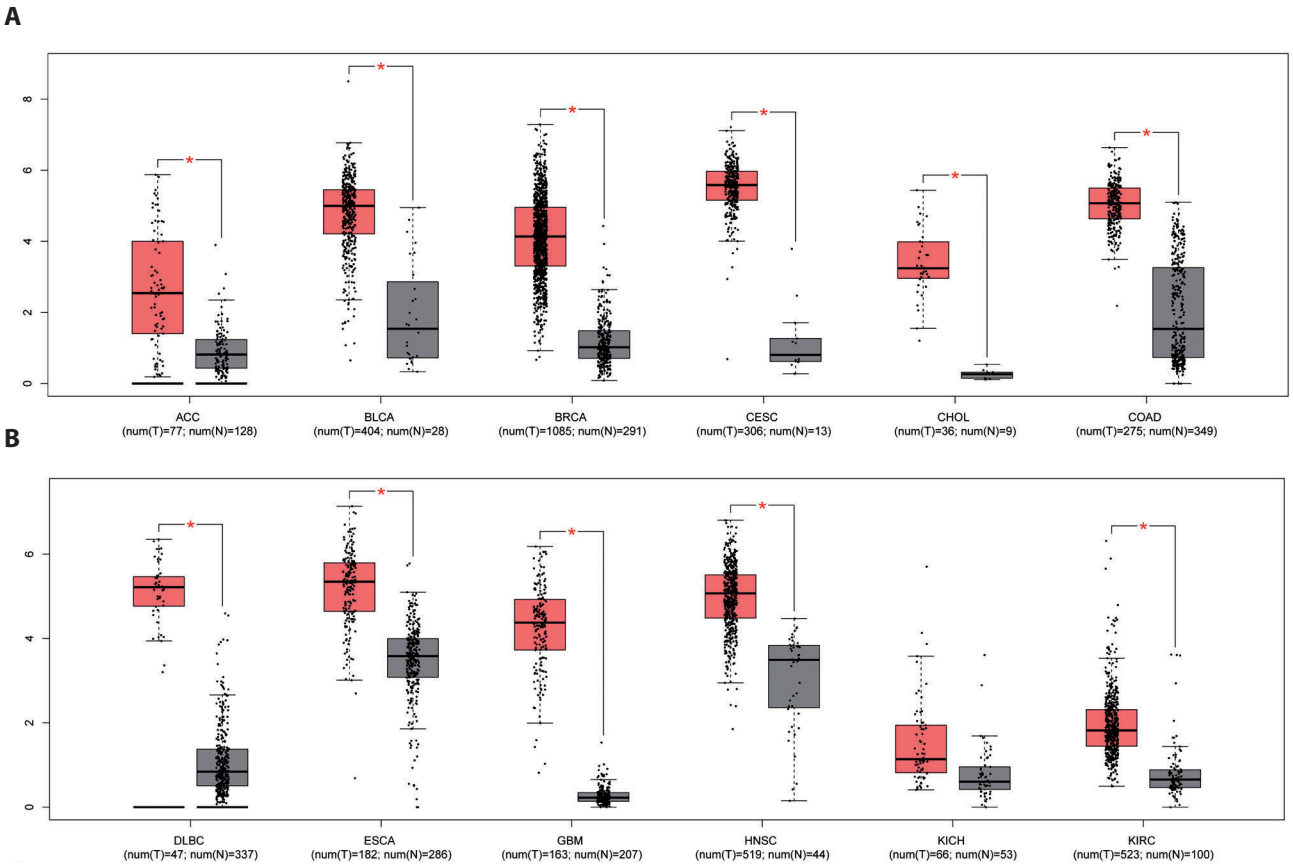


Figure S1. Expression of hub genes in (pancreatic cancer) based on pancreatitis status. Expression of CCNA2 (**A**), RDC4 (**B**), MCM7 (**C**), MCM7 (**D**), MCM5 (**E**) and MAD2L1 (**F**). Red boxes, PC with pancreatitis; blue boxes, normal adjacent tissues; orange boxes, PC without pancreatitis.



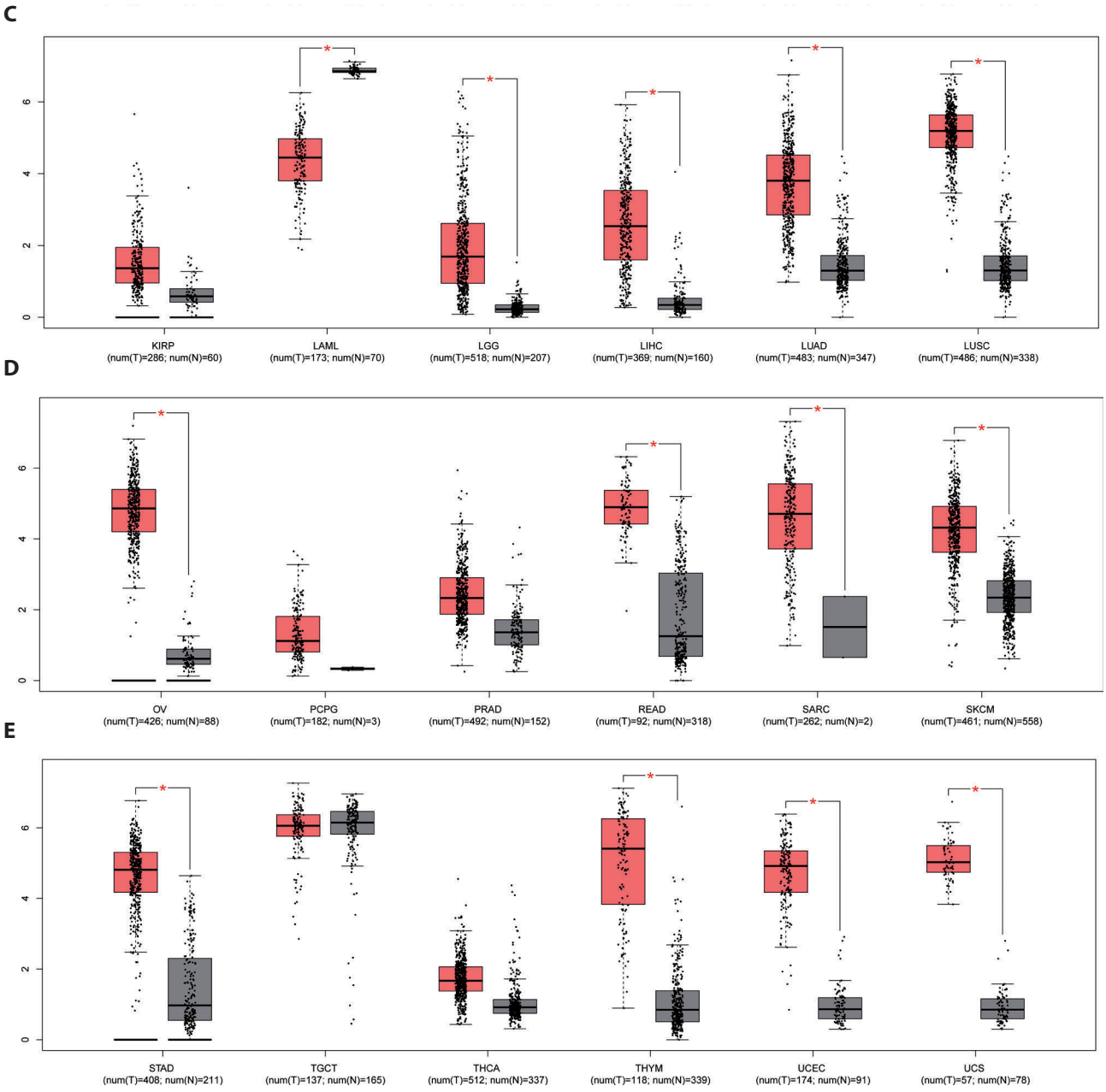
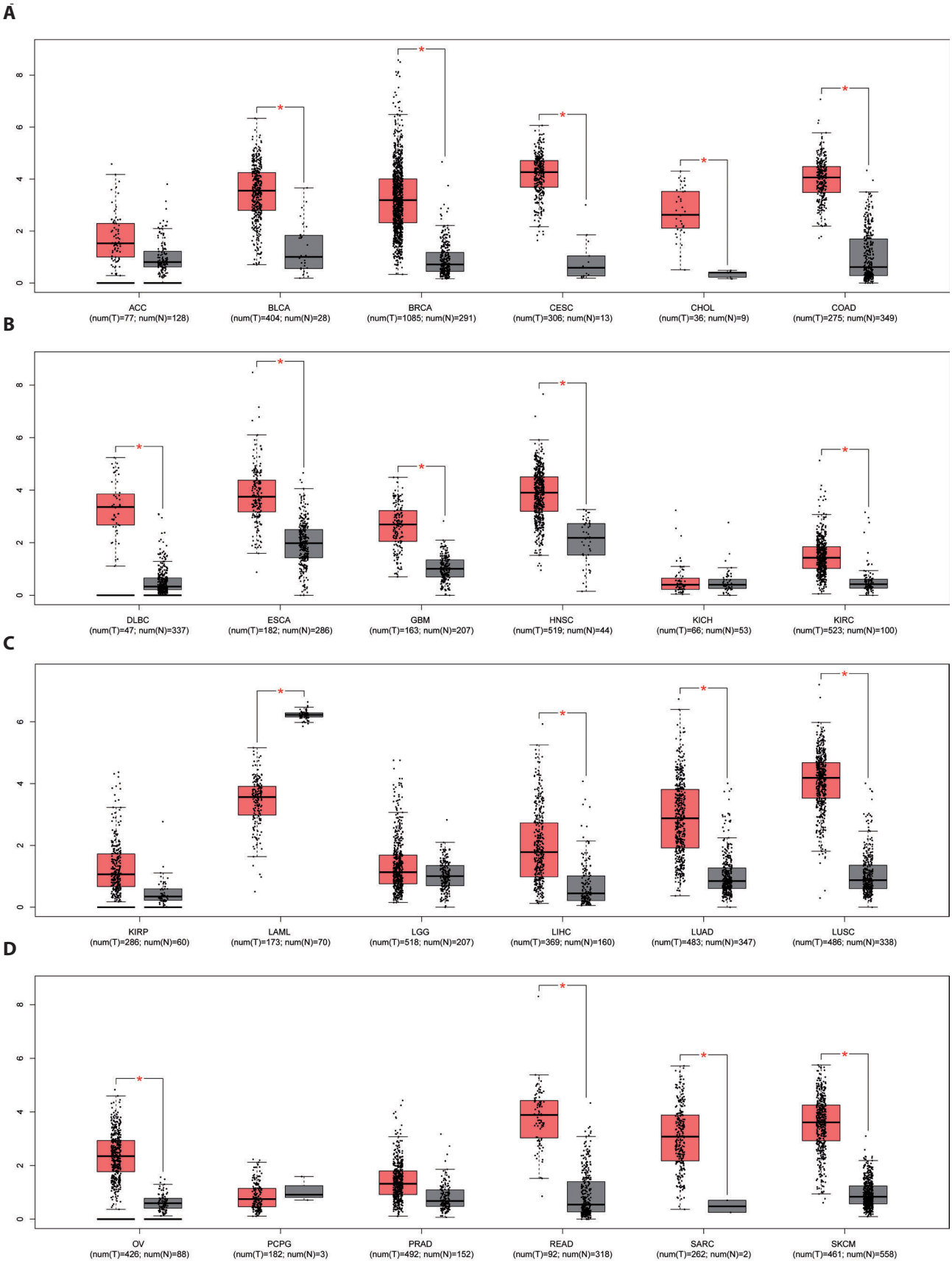


Figure S2. Pan-cancer analysis of CCNB2 expression (bar plot). Red boxes, tumor; gray boxes, normal adjacent tissues. Single-factor analysis of variance (ANOVA) was conducted to identify significant differences among the groups.


A
ACC
(num(T)=77; num(N)=128)
BLCA
(num(T)=404; num(N)=28)
BRCA
(num(T)=1085; num(N)=291)
CESC
(num(T)=306; num(N)=13)
CHOL
(num(T)=36; num(N)=9)
COAD
(num(T)=275; num(N)=349)
B
DLBC
(num(T)=47; num(N)=337)
ESCA
(num(T)=182; num(N)=286)
GBM
(num(T)=163; num(N)=207)
HNSC
(num(T)=519; num(N)=44)
KICH
(num(T)=66; num(N)=53)
KIRC
(num(T)=523; num(N)=100)
C
KIRP
(num(T)=286; num(N)=60)
LAML
(num(T)=173; num(N)=70)
LGG
(num(T)=518; num(N)=207)
LIHC
(num(T)=369; num(N)=160)
LUAD
(num(T)=483; num(N)=347)
LUSC
(num(T)=486; num(N)=338)
D
OV
(num(T)=426; num(N)=88)
PCPG
(num(T)=182; num(N)=3)
PRAD
(num(T)=492; num(N)=152)
READ
(num(T)=92; num(N)=318)
SARC
(num(T)=262; num(N)=2)
SKCM
(num(T)=461; num(N)=558)

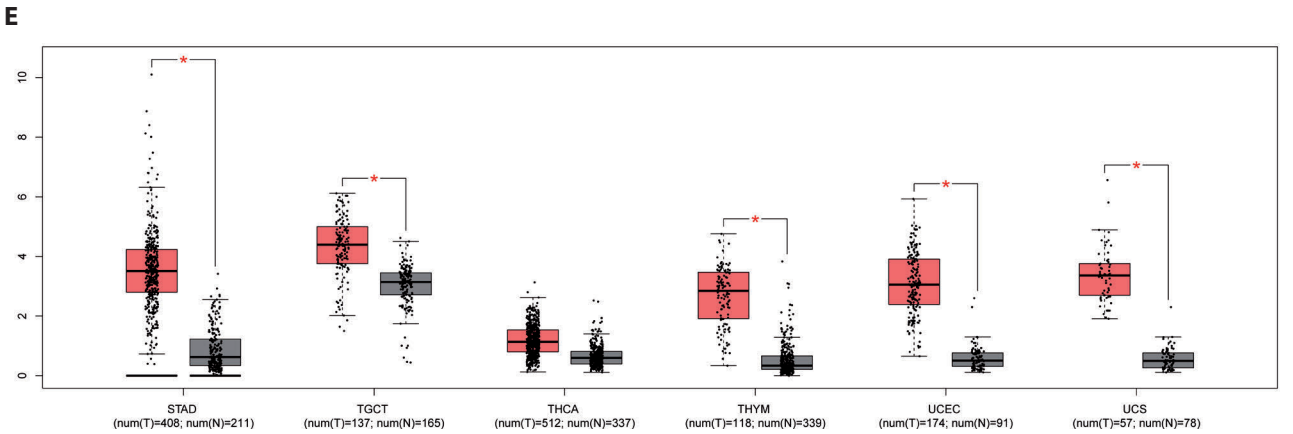
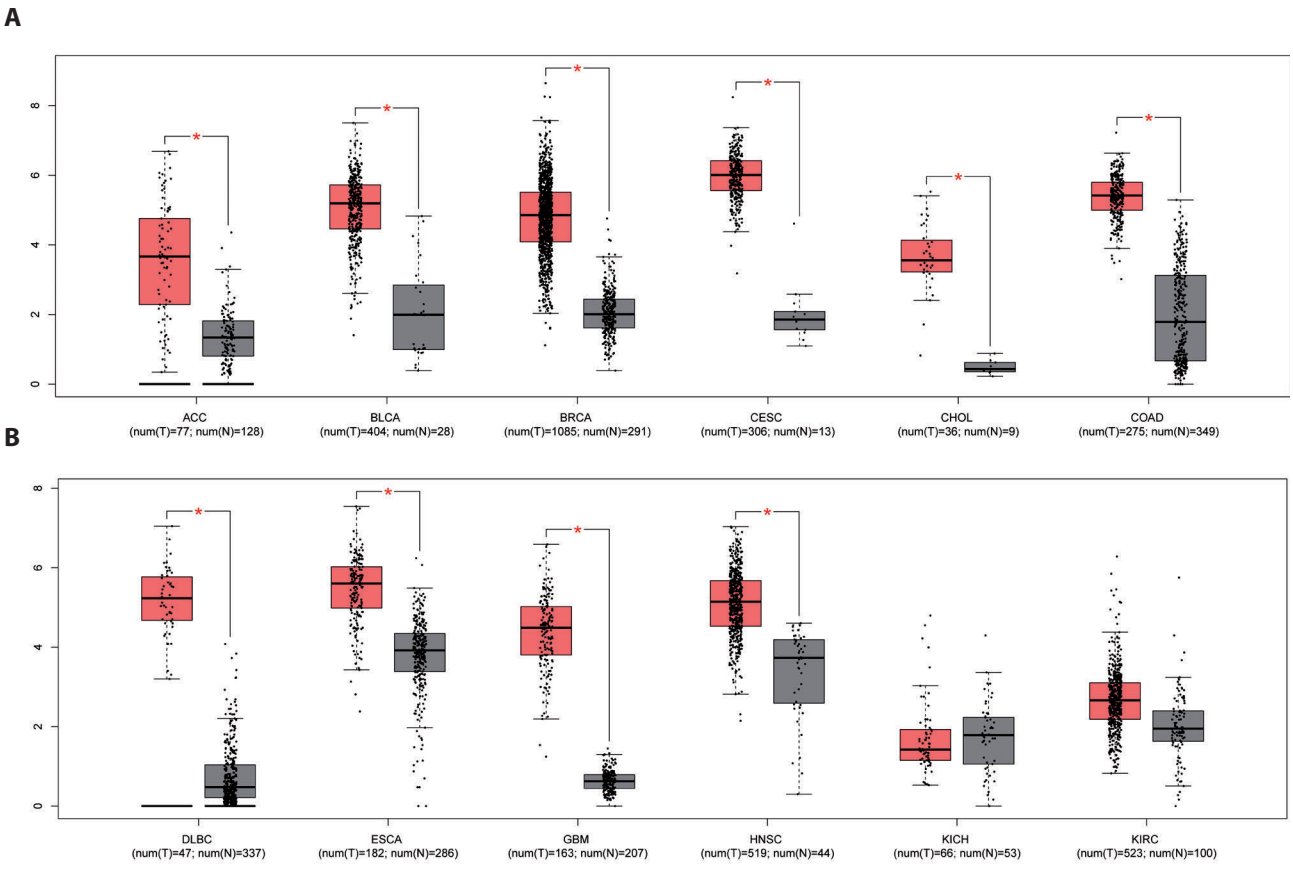


Figure S3. Pan-cancer analysis of CDC6 expression (bar plot). Red boxes, tumor; gray boxes, normal adjacent tissues. Single-factor analysis of variance (ANOVA) was conducted to identify significant differences among the groups.



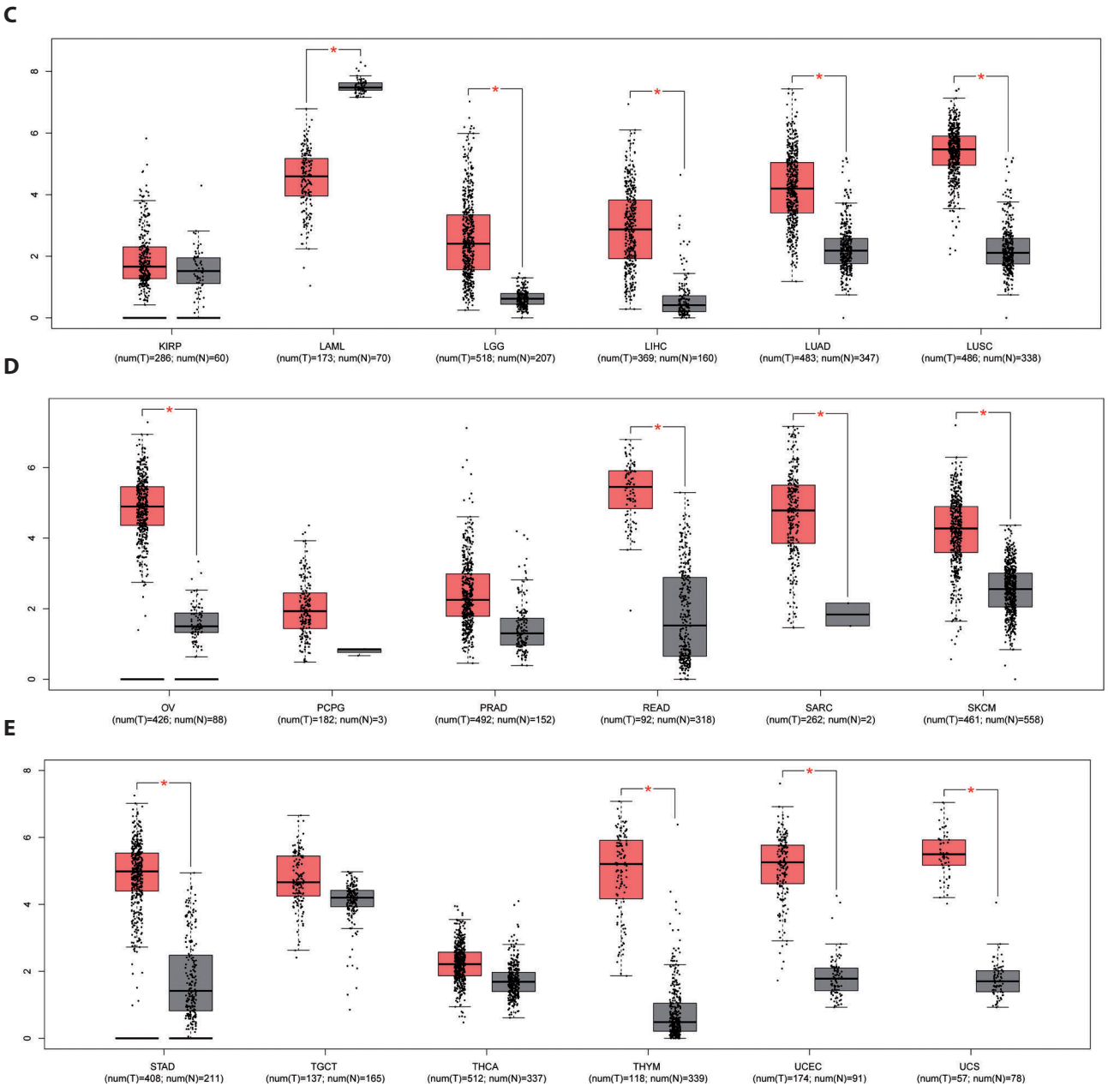
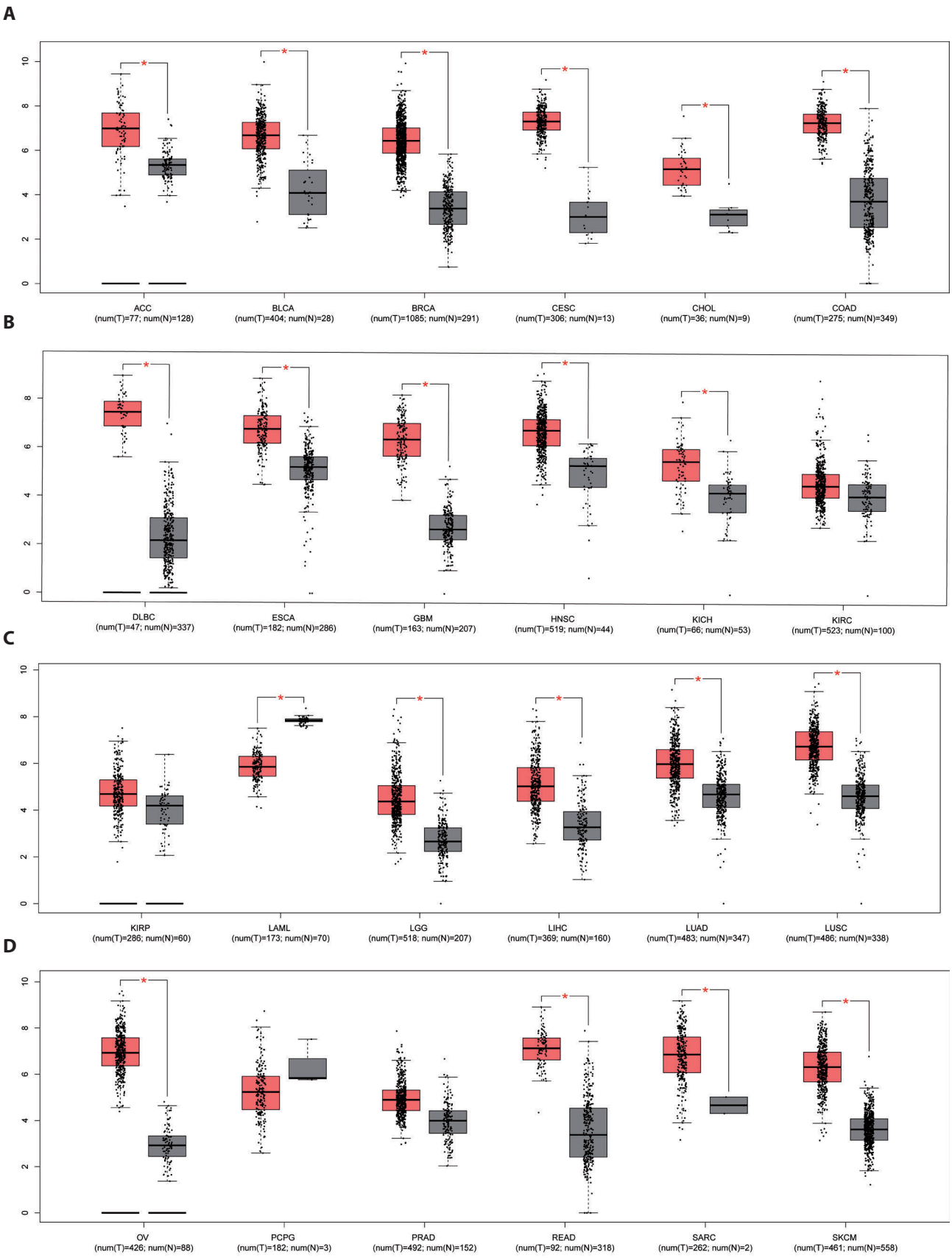


Figure S4. Pan-cancer analysis of CDK1 expression (bar plot). Red boxes, tumor; gray boxes, normal adjacent tissues. Single-factor analysis of variance (ANOVA) was conducted to identify significant differences among the groups.


A
ACC
(num(T)=77; num(N)=128)
BLCA
(num(T)=404; num(N)=28)
BRCA
(num(T)=1085; num(N)=291)
CESC
(num(T)=306; num(N)=13)
CHOL
(num(T)=36; num(N)=9)
COAD
(num(T)=275; num(N)=349)
B
DLBC
(num(T)=47; num(N)=337)
ESCA
(num(T)=182; num(N)=286)
GBM
(num(T)=163; num(N)=207)
HNSC
(num(T)=519; num(N)=44)
KICH
(num(T)=66; num(N)=53)
KIRC
(num(T)=523; num(N)=100)
C
KIRP
(num(T)=286; num(N)=60)
LAML
(num(T)=173; num(N)=70)
LGG
(num(T)=518; num(N)=207)
LIHC
(num(T)=369; num(N)=160)
LUAD
(num(T)=483; num(N)=347)
LUSC
(num(T)=486; num(N)=338)
D
OV
(num(T)=426; num(N)=88)
PCPG
(num(T)=182; num(N)=3)
PRAD
(num(T)=492; num(N)=152)
READ
(num(T)=92; num(N)=318)
SARC
(num(T)=262; num(N)=2)
SKCM
(num(T)=461; num(N)=558)

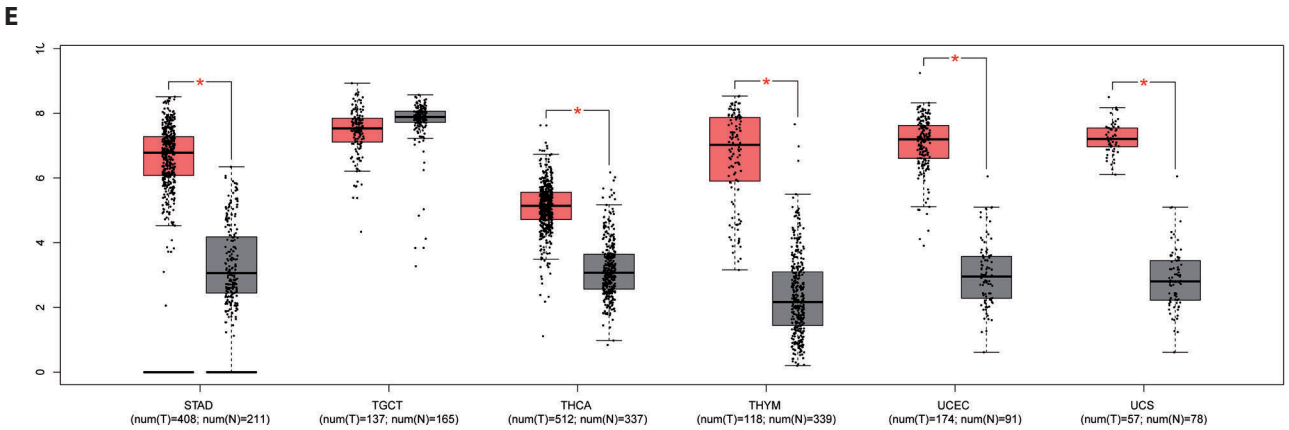


Figure S5. Pan-cancer analysis of CSK2 expression (bar plot). Red boxes, tumor; gray boxes, normal adjacent tissues. Single-factor analysis of variance (ANOVA) was conducted to identify significant differences among the groups.

Table S1. Full name of cancer-type abbreviations

Abbreviation	Full names
ACC	Adrenocortical carcinoma
BLCA	Bladder Urothelial Carcinoma
BRCA	Breast invasive carcinoma
CESC	Cervical squamous cell carcinoma and endocervical adenocarcinoma
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
OV	Ovarian serous cystadenocarcinoma
PCPG	Pheochromocytoma and Paraganglioma
PRAD	Prostate adenocarcinoma
READ	Rectum adenocarcinoma
SARC	Sarcoma
SKCM	Skin Cutaneous Melanoma
STAD	Stomach adenocarcinoma
TGCT	Testicular Germ Cell Tumors
THCA	Thyroid carcinoma
THYM	Thymoma
UCEC	Uterine Corpus Endometrial Carcinoma
UCS	Uterine Carcinosarcoma