Clinical Trial/Experimental Study

Medicine®

OPEN

A comprehensive investigation on pan-cancer impacts of constitutive centromere associated network gene family by integrating multi-omics data

A CONSORT-compliant article

Huimei Su, MS[a], Yuchun Fan, MS[a], Zhuan Wang, MS[a], Lihe Jiang, PhD[a,b,c,*]

Abstract

Background: The constitutive centromere associated network (CCAN) complex played a critical role in connecting the centromere with the mitotic spindle during mitosis and meiosis. Many studies have indicated that CCAN is related to the tumorigenesis and cancer development. Nonetheless, the overview of CCAN gene family in pan-cancer remain incompletely understood.

Methods: We performed a comprehensive investigation on pan-cancer impacts of CCAN by integrating multi-omics data. We comprehensively investigated the expression profile, kyoto encyclopedia of genes and genomes (kegg) pathway, mutation, copy number variation, tumor microenvironment, immune cells infiltration, and drug sensitivity of CCAN in pan-cancer. MRNA expression profiles were collected from the cancer genome atlas, oncomine and ccle, the differential expression and various relevance analysis were performed with R or Perl.

Results: The results showed that the expression of CCAN was different in 33 tumors. Intriguingly, the poor survival in adrenocortical carcinoma, cholangiocarcinoma, kidney chromophobe, mesothelioma, kidney renal clear cell carcinoma, brain lower grade glioma, pheochromocytoma and paraganglioma, prostate adenocarcinoma, thyroid carcinoma, uveal melanoma was most likely related to the kegg single transduction pathway including one carbon pool by folate, proteasome, arachidonic acid metabolism and so on. CENPC, ITGB3BP, APITD1, CENPU, and CENPW were more involved in tumor microenvironment, which more likely related to NK cells resting, T cells follicular helper, T cells CD8, neutrophils, macrophages M0, T cells CD4 memory activated. The relationship of CCAN expression with drug sensitivity showed that chelerythrine, nelarabine, and hydroxyurea maybe be potential drugs.

Conclusions: This multidimensional study provides a valuable resource to assist mechanism research and clinical utility about CCAN.

Abbreviations: CCAN = constitutive centromere associated network, ccle= broad institute cancer cell line encyclopedia, GSVA = gene set variation analysis, HR = hazard ratio, kegg = kyoto encyclopedia of genes and genomes, PCC = pearson correlation coefficient, TCGA = the cancer genome atlas.

Keywords: bioinformatics, constitutive centromere associated network, pan-cancer, TCGA

Editor: Sorush Niknamian.

This manuscript is a unique submission and is not being considered for publication by any other source in any medium. Furthermore, the manuscript has not been published, in part or in full, in any form.

The datasets generated and analyzed during the current study are available in the [UCSC xena] repository, [https://xenabrowser.net/datapages/], the [oncomine] repository, [https://www.oncomine.org/resource/login.html], the [human protein atlas data portal] repository, [https://www.proteinatlas.org/], the [ccle] repository, [Broad Institute Cancer Cell Line Encyclopedia (CCLE)].

The authors have no conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

[a] Medical College, Guangxi University, Nanning, Guangxi, P.R. China, [b] School of Basic Medical Sciences, Youjiang Medical University for Nationalities, Baise, Guangxi, P.R. China, [c] Key Laboratory of Translational Cancer Research, Putian University, Putian, Fujian, P.R. China.

[*] Correspondence: Lihe Jiang, Medical College, Guangxi University, Nanning 530004, School of Basic Medical Sciences, Youjiang Medical University for Nationalities, Baise, 533000, Guangxi, China (e-mail: jianglihe@ymun.edu.cn).

How to cite this article: Su H, Fan Y, Wang Z, Jiang L. A comprehensive investigation on pan-cancer impacts of constitutive centromere associated network gene family by integrating multi-omics data: A CONSORT-compliant article. Medicine 2022;00(00).Medicine 2022;101:7(e28821).

Received: 17 August 2021 / Received in final form: 25 January 2022 / Accepted: 26 January 2022

http://dx.doi.org/10.1097/MD.0000000000028821

1. Introduction

Precise chromosome segregation during mitosis and meiosis is indispensable to perpetuate genomic stability and prevent aneuploidy in eukaryotes.[1] Constitutive centromere associated network (CCAN) plays a pivotal role in precise chromosome segregation, which is comprised of *CENP-A/C/H/I/K/L/M/N/O/P/Q/R/S/T/U/W/X*[2]and required for chromosome congression, proper spindle attachment, mitotic checkpoint activity and mitosis sister chromatids separation, leading to the assembly of a functional kinetochore.[3] CCAN can be divided into two categories: one is monomer protein including CENPA and CENPC; the others are four multisubunit complexes including CENPL-CENPN, CENPH-CENPI-CENPK-CENPM, CENPO-CENPP-CENPQ-CENPR-CENPU, and CENPT-CENPW-CENPS-CENPX. During mitosis, CCAN, as an outer kinetochore platform, was recruited into the Y-shaped revelation of the CCAN, which was mediated by the DNA binding sulcus of the CENPL-CENPN subcomplex, to make CCAN subunits associate with nucleosomal DNA and ambient histone subunits, the CENP-A nucleosome was assembled to maintain the kinetochore structure and the stability of microtubule dynamics.[1] The information of CCAN was sketched in Table 1.

Early studies have confirmed that CCAN was dysregulation in many kinds of tumors, then changed the process and development of tumorigenesis, finally lead to poor or good prognosis. For instance, CENPA, CENPH, and CENPN were highly upregulated in hepatocellular carcinoma,[4–6] colorectal cancer,[7] breast cancer,[8,9] and promoted their progression. On the contrary, reduced expression of CENPE contributes to the development of hepatocellular carcinoma.[10] Interestingly, a report in 2017 has found that CENPR as a bilateral factor inhibits tumor in early stage but promotes tumor in late stage.[11] These results suggest that the abnormal expression of CCAN may closely relate to various stages of various tumors. The current cancer research of CCAN is not complete, such as the differential expression of CCAN in various cancers has not been clarified, and the regulatory mechanism behind the change of CCAN function has also not been deeply explored. Therefore, the research still needs to be deepened.

Some cancers in different organs have strong molecular similarities, while some cancers take up in the coequal organ but the molecular subtypes are completely different, on the contrary, they are more closely related to histological or anatomical types.[12] These findings supply us a new idea to search the common molecular characteristics of various human cancer through systematic analysis of specific genes in many kinds of cancers. This will provide a targeted basis for clinical comprehensive cancer diagnosis and precision medical treatment in the future.

In this study, we were aimed to explore the expression profile of CCAN, elaborate the signal transduction pathway, mutation, and copy number variation and the immunological characteristics, drug sensitivity. Our analysis attempt to systematize and understand the potential function of CCAN in cancer, and might help in the identification of novel markers providing several valid and testable hypotheses for the next exploration in cancer biology.

Table 1

Basic characteristics of the constitutive centromere associated network gene family.

Approved symbol	Approved name	Previous symbols	Aliases	Chromosome
CENPA	Centromere protein A		CENP-A, CenH3	2p23.3
CENPC	Centromere protein C	CENPC1	CENP-C, hcp-4, MIF2	4q13.2
CENPH	Centromere protein H			5q13.2
CENPI	Centromere protein I	FSHPRH1	LRPR1, CENP-I, Mis6	Xq22.1
CENPK	Centromere protein K		FKSG14, SOLT, CENP-K	5q12.3
CENPL	Centromere protein L	C1orf155	dJ383J4.3, FLJ31044	1q25.1
CENPM	Centromere protein M	C22orf18	Pane1, CENP-M, MGC861	22q13.2
CENPN	Centromere protein N	C16orf60	FLJ13607, FLJ22660, BM039	16q23.2
CENPO	Centromere protein O		MGC11266, CENP-O	2p23.3
CENPP	Centromere protein P		RP11-19J3.3, CENP-P	9q22.31
CENPQ	Centromere protein Q	C6orf139	FLJ10545, CENP-Q	6p12.3
ITGB3BP	Integrin subunit beta 3 binding protein		NRIF3, HSU37139, TAP20, CENPR	1p31.3
CENPS	Centromere protein S	APITD1, MHF1	CENP-S, FAAP16	1p36.22
CENPT	Centromere protein T	C16orf56	FLJ13111, CENP-T	16q22.1
CENPU	Centromere protein U	MLF1IP	CENP-U, KLIP1, CENP-50, PBIP1	4q35.1
CENPW	Centromere protein W	C6orf173	CUG2	6q22.32
CENPX	Centromere protein X	STRA13, MHF2	MGC14480, FAAP10, CENP-X	17q25.3

APITD1 = apoptosis-inducing TAF9-like domain 1, BM039 = centromere protein N, C16orf56 = centromere protein T, C16orf60 = centromere protein N, C22orf18 = centromere protein M, C6orf139 = centromere protein Q, C6orf173 = centromere protein W, CenH3 = histone H3-like centromeric protein, CENP-50 = centromere protein U, CENPA = centromere protein A, CENP-A = centromere protein A, CENPC = centromere protein C, CENP-C = centromere protein C, CENPC1 = centromere protein C 1, CENPH = centromere protein H, CENPI = centromere protein I, CENP-I = centromere protein I, CENPK = centromere protein K, CENP-K = centromere protein K, CENPL = centromere protein L, CENPM = centromere protein M, CENP-M = centromere protein M, CENPN = centromere protein N, CENPO = centromere protein O, CENP-O = centromere protein O, CENPP = centromere protein P, CENP-P = centromere protein P, CENPQ = centromere protein Q, CENP-Q = centromere protein Q, CENPR = integrin subunit beta 3 binding protein, CENPS = centromere protein S, CENP-S = centromere protein S, CENPT = centromere protein T, CENP-T = centromere protein T, CENPU = centromere protein U, CENP-U = centromere protein U, CENPW = centromere protein W, CENPX = centromere protein X, CENP-X = centromere protein X, C1orf155 = centromere protein L, CUG2 = centromere protein W, dJ383J4.3 = centromere protein L, FAAP10 = centromere protein X, FAAP16 = centromere protein S, FKSG14 = centromere protein K, FLJ10545 = centromere protein Q, FLJ13111 = centromere protein T, FLJ13607 = centromere protein N, FLJ22660 = centromere protein N, FLJ31044 = centromere protein L, FSHPRH1 = centromere protein I, hcp-4 = HoloCentric chromosome binding Protein, HSU37139 = integrin subunit beta 3 binding protein, ITGB3BP = integrin subunit beta 3 binding protein, KLIP1 = natural killer cell-specific antigen KLIP-1, LRPR1 = centromere protein I, MGC11266 = centromere protein O, MGC14480 = centromere protein X, MGC861 = centromere protein M, MHF1 = Mhf1p, MHF2 = Mhf2p, MIF2 = mini zinc finger 2, Mis6 = centromere connector protein mis 6, MLF1IP = centromere protein U, NRIF3 = integrin subunit beta 3 binding protein, Pane1 = proliferation associated nuclear element, PBIP1 = centromere protein U, RP11-19J3.3 = centromere protein P, SOLT = centromere protein K, STRA13 = stimulated by retinoic acid 13, TAP20 = integrin subunit beta 3 binding protein.

2. Materials and methods

2.1. Collection of expression profiles and other data

Thirty-three different kinds of human tumor data and survival data were downloaded from genomic data commons the cancer genome atlas (TCGA) cancer gene expression RNAseq-FPKM by UCSC xena (https://xenabrowser.net/datapages/), the download gene expression cohort information can be found in Table 2.

Twenty kinds of patient tumor tissue protein expression data between tumor and corresponding normal tissue were downloaded from the human protein atlas (https://www.proteinatlas.org/) chapter of pathology atlas. The tumor cellular expression of 16 genes (CENPA, CENPC, CENPH, CENPI, CENPK, CENPL, CENPM, CENPN, CENPO, CENPP, CENPQ, ITGB3BP (also known as CENPR), CENPT, MLF1IP (also known as CENPU), CENPW, STRA13 (also known as CENPX) were recruited from ccle (Broad Institute Cancer Cell Line Encyclopedia [CCLE]). The data of mutation or copy number was download from TCGA MC3 gene-level non-silent mutation or TCGA gistic2 thresholded by UCS xena.

2.2. Examination of the expression array of CCAN in pan carcinoma

Seventeen genes of UCSC xena were extracted for analysis: *CENPA*, *CENPC*, *CENPH*, *CENPI*, *CENPK*, *CENPL*, *CENPM*, *CENPN*, *CENPO*, *CENPP*, *CENPQ*, *ITGB3BP* (also known as *CENPR*), *APITD1* (also known as *CENPS*), *CENPT*, *CENPU*, *CENPW*, and *STRA13* (also known as *CENPX*). Ggpubr package in R was used to distinguish the differential expression of CCAN in 33 kinds of tumors then display in the form of heatmap and box plot. Besides, oncomine (https://www.oncomine.org/resource/login.html) is a powerful cancer microarray multifunctional database, we can search and visualize data rely on the standardized data. Seventeen genes were extracted for analysis: *CENPA*, *CENPC*, *CENPH*, *CENPI*, *CENPK*, *CENPL*, *CENPM*, *CENPN*, *CENPO*, *CENPP*, *CENPQ*, *ITGB3BP* (also known as *CENPR*), *CENPS*, *CENPT*, *CENPU*, *CENPW*, *STRA13* (also known as *CENPX*), then synthesized the pictures using microsoft office powerpoint. The high/medium expression patient ratio data was extracted protein expression summary

Table 2

Basic information of 33 kinds of cancers from UCSC Xena.

Version	Cancer type	Cancer name	Cancer cases	Normal cases
07-18-2019	ACC	Adrenocortical carcinoma	79	0
07-18-2019	BLCA	Bladder urothelial carcinoma	411	19
07-18-2019	BRCA	Breast invasive carcinoma	1104	113
07-19-2019	CESC	Cervical squamous cell carcinoma and endocervical adenocarcinoma	306	3
07-19-2019	CHOL	Cholangiocarcinoma	36	9
07-19-2019	COAD	Colon adenocarcinoma	471	41
07-19-2019	DLBC	Lymphoid neoplasm diffuse large B-cell Lymphoma	48	0
07-19-2019	ESCA	Esophageal carcinoma	162	11
07-19-2019	GBM	Glioblastoma multiforme	168	5
07-19-2019	HNSC	Head and neck squamous cell carcinoma	502	44
07-19-2019	KICH	Kidney chromophobe	65	24
07-19-2019	KIRC	Kidney renal clear cell carcinoma	535	72
07-19-2019	KIRP	Kidney renal papillary cell carcinoma	289	32
07-19-2019	LAML	Acute myeloid leukemia	151	0
07-19-2019	LGG	Brain lower grade glioma	529	0
07-19-2019	LIHC	Liver hepatocellular carcinoma	374	50
07-20-2019	LUAD	Lung adenocarcinoma	526	59
07-20-2019	LUSC	Lung squamous cell carcinoma	501	49
07-20-2019	MESO	Mesothelioma	86	0
07-20-2019	OV	Ovarian serous cystadenocarcinoma	379	0
07-20-2019	PAAD	Pancreatic adenocarcinoma	178	4
07-20-2019	PCPG	Pheochromocytoma and paraganglioma	183	3
07-20-2019	PRAD	Prostate adenocarcinoma	499	52
07-20-2019	READ	Rectum adenocarcinoma	167	10
07-20-2019	SARC	Sarcoma	263	2
07-20-2019	SKCM	Skin cutaneous melanoma	471	1
07-20-2019	STAD	Stomach adenocarcinoma	375	32
07-20-2019	TGCT	Testicular germ cell tumors	156	0
07-20-2019	THCA	Thyroid carcinoma	510	58
07-21-2019	THYM	Thymoma	119	2
07-21-2019	UCEC	Uterine corpus endometrial carcinoma	548	35
07-21-2019	UCS	Uterine carcinosarcoma	56	0
07-21-2019	UVM	Uveal melanoma	80	0

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, 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, TGCT = Testicular germ cell tumors, THCA = Thyroid carcinoma, THYM = Thymoma, UCEC = Uterine corpus endometrial carcinoma, UCS = Uterine carcinosarcoma, UVM = Uveal melanoma.

from pathology atlas of the human protein atlas, calculate the high/medium expression patient ratio by microsoft office excel and draw rose diagram by R. The CCAN expression profile in 6 types of cell line was showed with heatmap by R. There are only 8 genes were recruited for analysis: *CENPA*, *CENPH*, *CENPL*, *CENPM*, *CENPP*, *CENPQ*, *ITGB3BP* (also known as *CENPR*), *CENPT*.

2.3. Examination of CCAN expression with risk factors for survival in pan carcinoma

Survival package in R was utilized to count hazard ratio (HR) of UCSC xena tumor gene expression data, the HR was adjusted at log as base 10 logarithm, then draw forest map to show HR fluctuation range.

2.4. Examination of CCAN expression related signal transduction pathway in pan carcinoma

The gene set variation analysis (GSVA) (http://www.gsea-msigdb.org/gsea/index.jsp) is one of the gene set enrichment analysis methods which can evaluate the variation of pathway activity in a simple population with an unsupervised way. In this paper, we elected MsigDB *KEGG* gene sets in gene set enrichment analysis to analyze. GSVA and gseabase packages in R were utilized to change the amount of gene expression to the amount of signal pathway in the patient sample. The relationship between CCAN expression and pathway activity was calculated by Pearson correlation coefficient (PCC). We selected the CCAN intersection pathway with $P < .05$ to draw the correlation coefficient table by R, and selected the pathway with $P < .05$ and $|\text{Cor}| > 0.3$ to draw the histogram of relate pathway enrich number by R and Perl.

2.5. Examination of CCAN expression with mutation in pan carcinoma

We used Microsoft office excel to calculate the frequency of mutation then display in the form of heatmap by R.

2.6. Examination of CCAN expression with copy number variation in pan carcinoma

We used Microsoft office excel to calculate the frequency of copy number variation, reshape2 and rcolorbrewer packages in R were utilized to draw the histogram. The upper left and lower right represent deletion and amplification, respectively.

2.7. Examination of CCAN expression with tumor microenvironment in pan carcinoma

Estimate package in R was utilized to change the amount of gene expression to the score of stromal cell or immune cell. Then the PCC was calculated and shown in the relevant picture.

2.8. Examination of CCAN expression with tumor immune cell infiltration in pan carcinoma

Cibersort (CIBERSORTx [stanford.edu]) is a tool to deconvolute the expression matrix of human immune cell subtypes. We calculated the PCC with $P < .05$ between CCAN expression and immune infiltration cell. Vegan, dplyr, and corrplot packages in R were utilized to draw the correlation heatmap.

2.9. Examination of CCAN expression with drug sensitivity in pan carcinoma

Cellminer (CellMiner—Analysis Tools [nih.gov]) is a drug screening tool that covers the genomic target information of thousands of drugs. Ggpubr package in R was utilized to gain gene related drugs then display in the form of linear correlation graph.

2.10. Ethical review

The databases involved in this study are open and can be used by researchers for unlimited times. Their source data have been published and approved by the local medical ethics committee. Therefore, there is no need to provide the approval documents of the ethics committee, and there are no moral problems and other conflicts of interest.

3. Results

3.1. CCAN expression profile at mRNA level

From the UCSC xena database of TGCA, we extracted 17 CCNA gene family members expression profiles from 33 tumors tissues sample, classified tumor group and normal group, and explored the differential expression. According to cluster results, CCAN was divided into three groups: CENPW, CENPH, CENPA, CENPU, CENPM were high expression group, they were strongly expressed in breast invasive carcinoma, uterine corpus endometrial carcinoma, lung adenocarcinoma, head and neck squamous cell carcinoma, stomach adenocarcinoma, liver hepatocellular carcinoma, bladder urothelial carcinoma, esophageal carcinoma, cholangiocarcinoma, glioblastoma multiforme, lung squamous cell carcinoma. CENPL, CENPI, CENPO, CENPK, CENPN, and STRA13 were medium expression group, they were slightly highly expressed in those tumor tissues. APITD1, CENPP, ITGB3BP, CENPQ, CENPT, and CENPC were low expression group, they were lowly expressed in almost all tumor tissues (Fig. 1A). After excluding the cancer types of <5 corresponding normal samples, the specific distinctive expression of CCAN gene in other cancers were visualized in Figure 1B.

To verify our TGCA finding, we examined the expression of CCAN from the perspective of oncomine. The results showed that except CENPC, CENPP, ITGB3BP, CENPS, and CENPT present almost medium expression level, others gene appear to highly expressed across the 20 kinds of cancers (Fig. 1C).

3.2. CCAN expression profile at protein and cell level

Limited by incomplete data of gene and cancer type, we used the count data by the protein atlas. The high/medium expression patient ratio of some CCAN was summarized in Figure 2A. We selected 6 kinds of tumors included in immunohistochemistry to observe the mRNA expression of CCAN in corresponding tumor cells. Preliminary results showed that STRA13, CENPH, CENPN, CENPA, CENPO, MLF1IP, and CENPW were highly expression in cancers, and the expression of CENPP, CENPC, ITGB3BP, CENPI, CENPL, CENPT, CENPK, and CENPQ were lower (Fig. 2B).

3.3. Analysis of the characteristics of CCAN

We used cox regression analysis to analyze the relationship between CCAN and prognosis in different cancers. Most genes

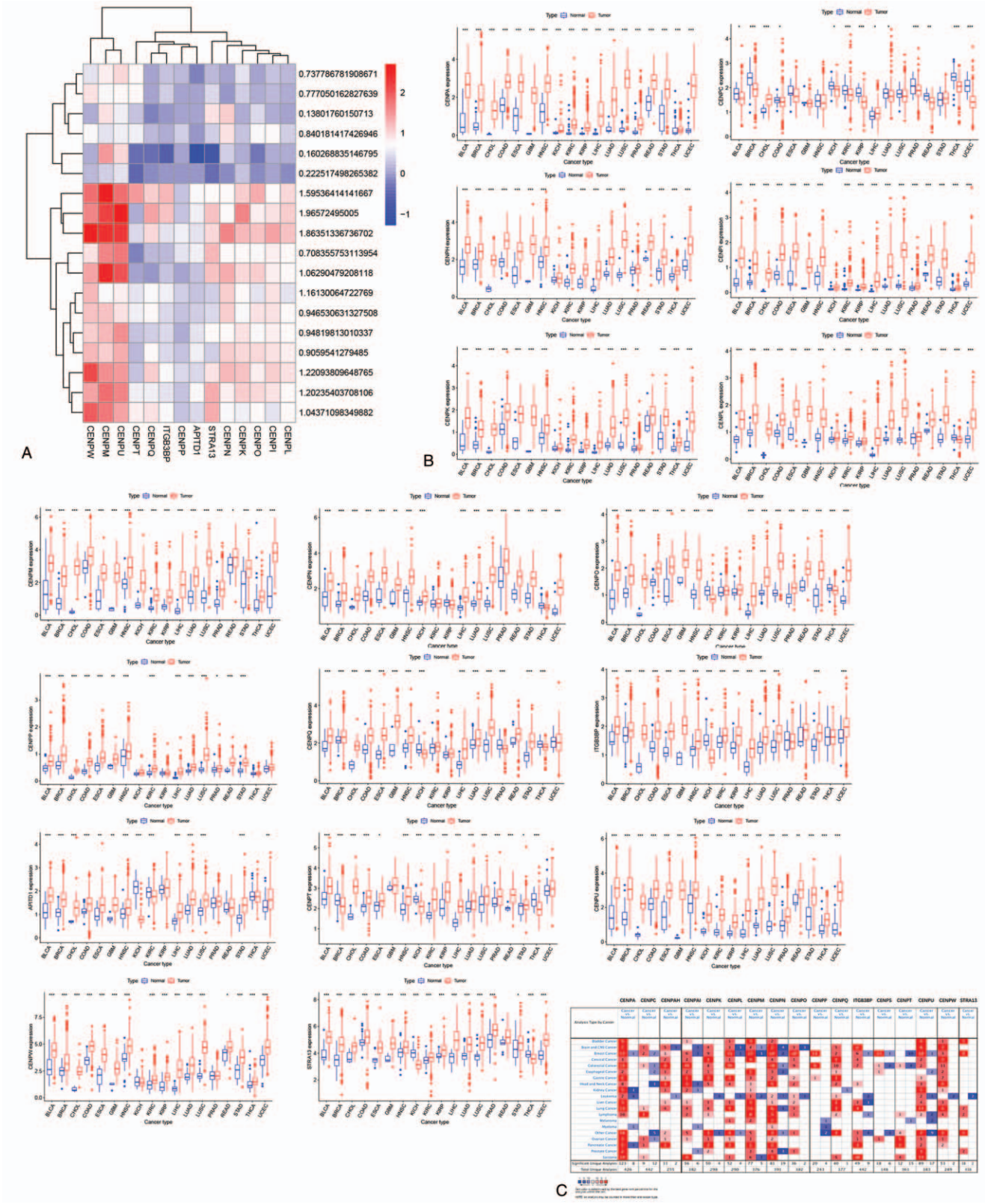

Figure 1. CCAN expression profile at mRNA level. (A) CCAN expression in different cancer and normal tissues of TCGA. The color represented the log2 fold change value, the expression decreased from red to blue. (B) CCAN expression in 17 types of cancers between cancer and normal tissues of TCGA. The "***", "**," "*" represent $P<.001$, $P<.01$, $P<.05$ individually. (C) CCAN expression in 20 types of cancer and normal tissues of Oncomine.

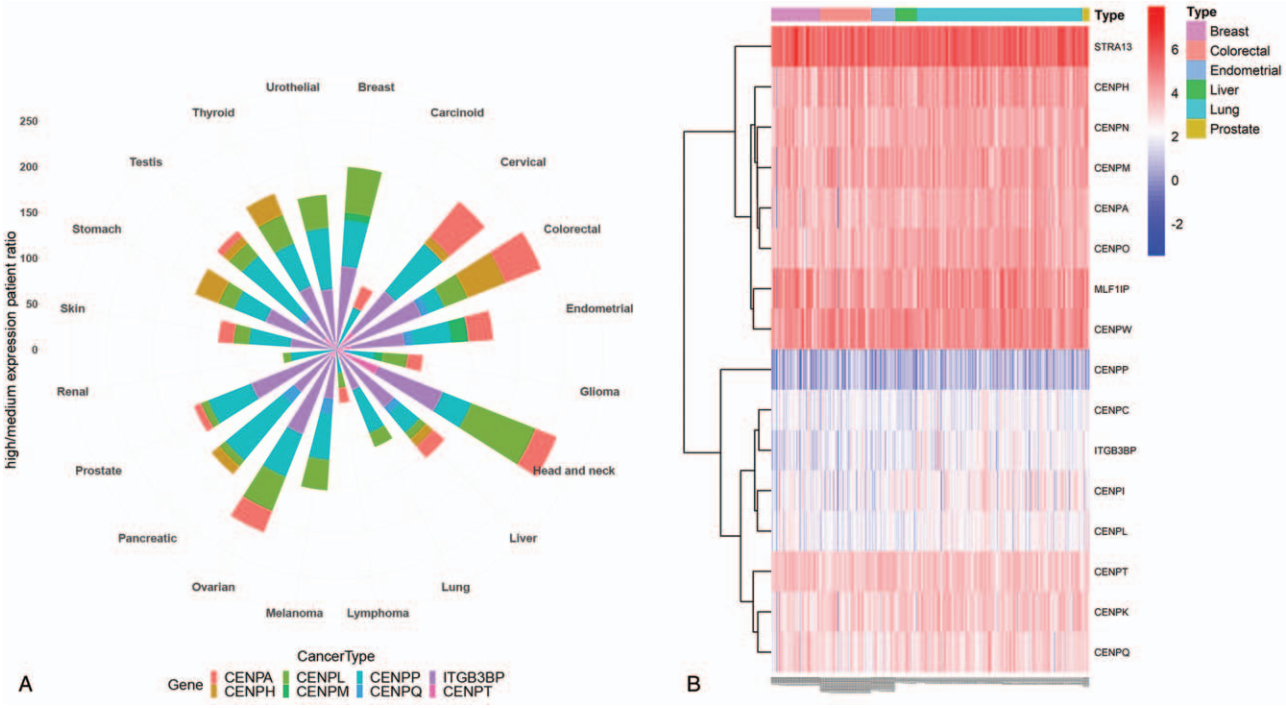

Figure 2. CCAN expression profile at protein and cell level. (A) Expression of CCAN protein in tumor tissues by high/medium expression patient ratio (%). (B) CCAN expression in different cancer cell line of CCLE.

are high risk genes in adrenocortical carcinoma, cholangiocarcinoma, kidney chromophobe, mesothelioma, brain lower grade glioma, kidney renal clear cell carcinoma, pheochromocytoma and paraganglioma, prostate adenocarcinoma, thyroid carcinoma, uveal melanoma, except for CENPC, CENPM, CENPP, CENPT, CENPU, CENPX (Fig. 3).

The relationship of kyoto encyclopedia of genes and genomes (kegg) single transduction pathways and CCAN were analyzed GSVA. We counted the number of each CCAN associated with activation and inhibition pathways, and found that CENPO and APITD1 were more likely to be interrelated to the positive pathway, whereas CENPT interrelated to the negative pathway (Fig. 4 A). Combined with all the CCAN gene members and significantly related pathways, a total of 74 intersection pathways were obtained (Table 3), including one carbon pool by folate, proteasome, cell cycle, arachidonic acid metabolism, metabolism of xenobiotics by cytochrome p450, linoleic acid metabolism, primary bile acid biosynthesis, pyrimidine metabolism, basal transcription factors, and so on.

Through the extraction and summary of all the samples of TCGA of UCSC xena, we counted the mutation frequency of CCAN in 33 kinds of cancer (Fig. 4B). From the results, CCAN rarely changed the overall mutation rate, the mutation was mainly concentrated in colon adenocarcinoma, rectum adenocarcinoma and uterine corpus endometrial carcinoma. Especially, CENPI and CENPC in uterine corpus endometrial carcinoma were more significant.

In addition, as for copy number variation of CCAN in 33 types of cancer, we counted homozygous deletion and single copy deletion to copy number deletion, and counted diploid normal

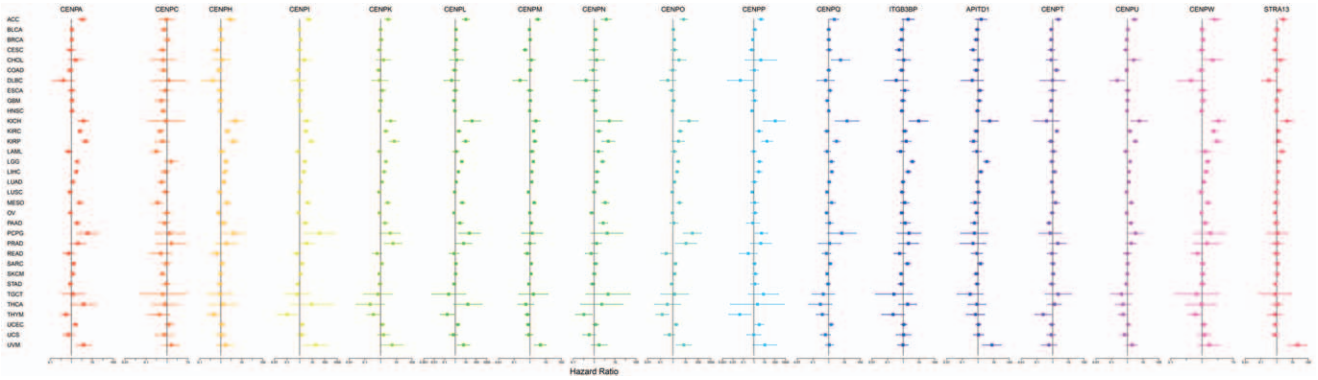

Figure 3. The cox analysis of CCAN in different cancer types.

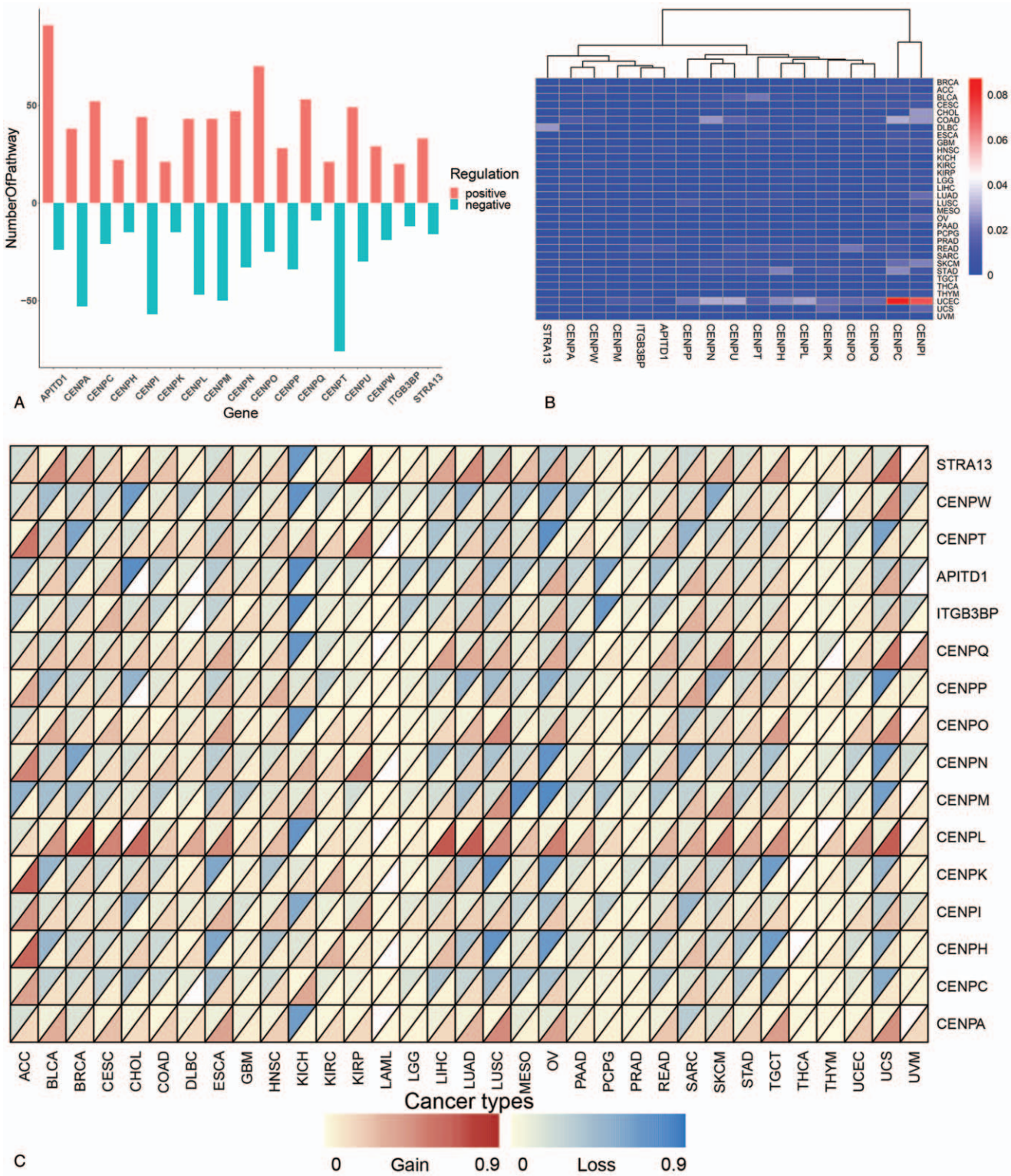

Figure 4. The characteristics around CCAN. (A) The number of correlated KEGG pathways in each individual CCAN. (B) Mutation frequency of CCAN in various cancers. (C) The copy number variations frequency of CCAN in various cancers.

copy and low-level copy number amplification to copy number amplification (Fig. 4C). On the whole, there were slight copy number changes in varying degrees, especially in CENPA, CENPL, CENPQ, CENPT, and STRA13 of amplification and CENPC, CENPH, CENPM, CENPN, CENPP, and CENPW of deletion.

3.4. Analysis of immune and drug correlation with CCAN

In order to understand the immune characteristics of CCAN, we first investigated the immune microenvironment. The outcomes showed that CENPC, CENPI, CENPP, CENPW were most positively related to the stromal cell, and CENPC, ITGB3BP,

Table 3

Correlation between the constitutive centromere associated network expression and signal transduction pathways.

	Correlation coefficient					
Pathway	**CENPA**	**CENPC**	**CENPH**	**CENPI**	**CENPK**	**CENPL**
Alanine aspartate and glutamate metabolism	−0.37	−0.09	0.1	−0.13	−0.01	0
Aldosterone regulated sodium reabsorption	−0.33	−0.25	−0.12	−0.13	−0.04	−0.45
Aminoacyl trna biosynthesis	0.02	0.15	0.02	0.04	0.01	0.05
Antigen processing and presentation	−0.24	−0.23	0.27	0.05	0.27	−0.38
Arachidonic acid metabolism	−0.34	−0.48	−0.09	−0.36	−0.2	−0.22
Arrhythmogenic right ventricular cardiomyopathy arvc	0.11	0.05	0.03	0.06	0.07	−0.02
Ascorbate and aldarate metabolism	−0.36	−0.09	−0.04	−0.06	0.01	−0.15
Axon guidance	0.09	0.28	0.01	−0.04	0	0.03
Basal cell carcinoma	−0.04	0.23	−0.02	−0.12	−0.04	−0.03
Basal transcription factors	0.25	0.08	0.29	0.34	0.24	0.3
Base excision repair	0.04	0.19	0.14	0.12	0.04	0.07
Butanoate metabolism	−0.23	0.19	−0.04	−0.18	0.02	−0.09
Cell adhesion molecules cams	−0.14	−0.31	0.19	−0.01	0.02	−0.2
Cell cycle	0.42	0.1	0.27	0.34	0.74	0.29
Chronic myeloid leukemia	0.02	0.02	0.05	0	0.01	0.17
Colorectal cancer	0.1	0.09	0.02	0.11	0.04	0.36
Dna replication	0.31	−0.16	0.18	0.17	0.11	0.14
Dorso ventral axis formation	0.02	0.07	0.01	0.04	0	−0.08
Drug metabolism cytochrome p450	−0.47	−0.18	−0.21	−0.44	−0.21	−0.28
Fatty acid metabolism	−0.13	0.11	−0.04	−0.11	−0.19	0.02
Gap junction	−0.02	0.1	0.02	−0.05	0.02	−0.07
Glutathione metabolism	−0.25	0.16	−0.11	−0.09	−0.03	0.01
Glycosphingolipid biosynthesis ganglio series	−0.01	−0.03	−0.16	−0.04	0.06	0.1
Glycosphingolipid biosynthesis lacto and neolacto series	−0.05	0.05	−0.17	−0.07	−0.14	−0.22
Gnrh signaling pathway	−0.04	−0.04	0.02	−0.09	−0.06	−0.08
Histidine metabolism	−0.2	0.22	−0.14	−0.21	−0.15	−0.19
Homologous recombination	0.12	0.05	0.13	0.12	0.09	0.16
Inositol phosphate metabolism	0.12	0.32	0.03	−0.11	0	0.01
Intestinal immune network for iga production	−0.03	−0.11	0.25	−0.22	0.08	−0.2
Leukocyte transendothelial migration	−0.03	−0.13	−0.15	0.07	0.04	−0.25
Linoleic acid metabolism	−0.28	−0.04	−0.22	−0.34	−0.19	−0.49
Long-term depression	−0.09	−0.2	−0.03	0	0.08	−0.25
Long-term potentiation	0.05	0.11	−0.02	0.07	0.05	−0.17
Lysine degradation	−0.15	0.08	−0.22	−0.11	−0.14	−0.03
Melanogenesis	0.04	0.31	0.01	−0.03	0.08	0.02
Melanoma	0.01	−0.02	−0.01	0.03	0.05	0
Metabolism of xenobiotics by cytochrome p450	−0.43	−0.23	−0.48	−0.43	−0.34	−0.27
Mismatch repair	0.11	0	0.16	0.08	0.06	0.14
Neuroactive ligand receptor interaction	−0.16	−0.34	−0.03	−0.16	−0.03	−0.12
Neurotrophin signaling pathway	−0.1	0.11	−0.02	−0.06	0.04	−0.25
Non homologous end joining	0.31	0.09	0.11	0.15	0.14	0.07
Nucleotide excision repair	0.14	0.07	0.16	0.16	0.14	0.24
Olfactory transduction	−0.11	−0.14	−0.13	−0.16	−0.16	−0.14
One carbon pool by folate	0.62	0.24	0.13	0.06	0.16	0.2
Oocyte meiosis	0.24	0.03	0.2	0.18	0.13	0.13
P53 signaling pathway	0.16	−0.09	0.15	0.19	0.18	0.15
Pancreatic cancer	0.1	−0.01	0.06	0.05	0.03	0.22
Pantothenate and coa biosynthesis	0.14	−0.03	0.13	0.09	0.13	0
Pathogenic escherichia coli infection	0.18	−0.14	0.13	0.18	0.13	0.05
Pathways in cancer	0.08	0.29	0.05	0.1	0.08	0.2
Primary bile acid biosynthesis	−0.5	−0.01	−0.24	−0.26	−0.02	−0.25
Primary immunodeficiency	−0.25	−0.22	0.14	−0.19	0.11	−0.11
Progesterone mediated oocyte maturation	0.13	0.02	0.31	0.12	0.1	0.18
Propanoate metabolism	−0.11	0.09	−0.15	−0.18	−0.19	−0.15
Proteasome	0.53	0.03	0.12	0.1	0.26	0.12
Pyrimidine metabolism	0.2	0.12	0.14	0.23	0.2	0.14
Regulation of actin cytoskeleton	0	−0.17	0	0.14	0.11	−0.08
Renin angiotensin system	0.01	−0.08	0.1	−0.07	0.08	−0.1
Riboflavin metabolism	0.21	−0.12	0.23	0.09	0.02	−0.17
Rna degradation	0.36	0.24	0.19	0.25	0.18	0.14
Rna polymerase	0.05	0.15	0.14	0.17	0.1	0.11

(*continued*)

Table 3

(continued).

Pathway	Correlation coefficient					
	CENPA	CENPC	CENPH	CENPI	CENPK	CENPL
Small cell lung cancer	0.09	0.36	0	0.11	0.04	0.25
Spliceosome	0.23	0.09	0.25	0.1	0.41	0.25
Steroid hormone biosynthesis	−0.33	−0.1	−0.03	−0.09	−0.39	−0.47
Thyroid cancer	0.38	−0.19	−0.09	0.33	0.03	0.16
Tryptophan metabolism	−0.28	−0.01	−0.08	−0.17	−0.22	−0.08
Type II diabetes mellitus	−0.28	0.03	−0.09	−0.08	0.06	−0.25
Type I diabetes mellitus	−0.33	−0.09	0.06	−0.23	0.2	−0.09
Tyrosine metabolism	−0.41	0.09	−0.14	−0.21	−0.27	−0.21
Ubiquitin mediated proteolysis	0.19	0.22	0.09	0.15	0.07	0.24
Valine leucine and isoleucine degradation	−0.1	0.23	−0.02	−0.12	−0.16	−0.04
Vasopressin regulated water reabsorption	0.2	0.04	0.11	0.08	0.1	−0.1
Viral myocarditis	0.11	−0.24	0.14	0.17	0.15	−0.34
Wnt signaling pathway	0.09	0.46	0.02	−0.01	0.08	0.13

Pathway	Correlation coefficient					
	CENPM	CENPN	CENPO	CENPP	CENPQ	ITGB3BP
Alanine aspartate and glutamate metabolism	−0.16	−0.11	−0.13	0.01	−0.12	0.12
Aldosterone regulated sodium reabsorption	−0.22	−0.09	−0.21	−0.51	−0.05	−0.12
Aminoacyl trna biosynthesis	0.19	0.05	0.09	0.15	0.01	−0.02
Antigen processing and presentation	0.2	−0.34	−0.16	−0.27	−0.16	0.24
Arachidonic acid metabolism	−0.28	−0.6	−0.47	−0.25	−0.08	−0.14
Arrhythmogenic right ventricular cardiomyopathy arvc	0.06	0.02	0.03	−0.08	0.04	−0.1
Ascorbate and aldarate metabolism	−0.18	−0.4	−0.11	−0.25	−0.12	−0.15
Axon guidance	0.02	0.01	0.17	−0.08	0.09	−0.08
Basal cell carcinoma	−0.08	−0.04	0.03	0.29	−0.02	−0.08
Basal transcription factors	0.3	0.59	0.32	0.22	0.25	−0.01
Base excision repair	0.15	0.15	0.1	0.07	0.18	−0.09
Butanoate metabolism	−0.23	−0.12	−0.11	−0.03	−0.11	0.02
Cell adhesion molecules cams	−0.04	−0.14	−0.23	−0.22	0.11	0.09
Cell cycle	0.35	0.34	0.43	0.21	0.25	0.05
Chronic myeloid leukemia	−0.11	0.07	0.07	−0.08	0.09	−0.07
Colorectal cancer	0.1	0.13	0.46	0.27	0.15	−0.08
Dna replication	0.29	0.34	0.15	0.22	0.24	0.03
Dorso ventral axis formation	−0.04	−0.07	0.14	0.13	0.02	−0.15
Drug metabolism cytochrome p450	−0.24	−0.41	−0.25	−0.22	−0.12	−0.14
Fatty acid metabolism	−0.35	0.03	0.11	0.09	−0.11	−0.04
Gap junction	−0.02	−0.05	0.1	−0.08	0.04	−0.2
Glutathione metabolism	−0.26	−0.17	−0.04	0.07	−0.13	−0.11
Glycosphingolipid biosynthesis ganglio series	−0.14	−0.04	0.12	0.08	−0.1	−0.16
Glycosphingolipid biosynthesis lacto and neolacto series	−0.14	0	−0.08	−0.14	0.12	−0.2
Gnrh signaling pathway	−0.11	−0.01	−0.13	−0.09	0.1	−0.08
Histidine metabolism	−0.2	−0.27	−0.15	0.02	−0.13	−0.12
Homologous recombination	0.19	0.18	0.18	0.11	0.13	0.03
Inositol phosphate metabolism	0.3	−0.08	−0.1	0.11	0.08	0
Intestinal immune network for iga production	−0.25	−0.01	−0.22	−0.2	0.3	0.18
Leukocyte transendothelial migration	−0.01	0.02	−0.17	−0.21	0.12	−0.05
Linoleic acid metabolism	−0.22	−0.52	−0.56	−0.16	−0.06	−0.16
Long term depression	0.06	0.03	0	−0.16	−0.07	−0.05
Long term potentiation	−0.1	0.06	0.02	−0.09	0.04	−0.08
Lysine degradation	−0.18	0.01	0.02	0.28	−0.04	−0.1
Melanogenesis	−0.04	0.05	−0.04	−0.1	−0.03	−0.09
Melanoma	−0.02	−0.02	0.02	−0.08	−0.01	−0.1
Metabolism of xenobiotics by cytochrome p450	−0.44	−0.63	−0.39	−0.19	−0.22	−0.09
Mismatch repair	0.16	0.17	0.14	0.23	0.24	−0.06
Neuroactive ligand receptor interaction	−0.21	−0.08	−0.22	−0.22	0.1	0.05
Neurotrophin signaling pathway	−0.09	0.11	0.04	−0.09	0.12	−0.13
Non homologous end joining	0.21	0.19	0.15	0.13	0.1	0.03
Nucleotide excision repair	0.24	0.33	0.23	0.21	0.24	0.02
Olfactory transduction	−0.3	−0.17	−0.08	−0.17	−0.1	−0.19
One carbon pool by folate	0.13	0.07	0.36	0.15	0.15	−0.04

(*continued*)

Table 3

(continued).

Pathway	Correlation coefficient					
	CENPM	CENPN	CENPO	CENPP	CENPQ	ITGB3BP
Oocyte meiosis	0.16	0.17	0.24	0.37	0.22	0
P53 signaling pathway	0.21	0.2	0.41	0.34	0.12	0.03
Pancreatic cancer	0.08	0.25	0.28	0.23	0.26	−0.05
Pantothenate and coa biosynthesis	0.14	0.11	0.14	0.17	0.19	0.06
Pathogenic escherichia coli infection	0.12	0.01	0.18	−0.13	0.29	0.1
Pathways in cancer	0.07	0.11	0.13	−0.08	0.08	−0.04
Primary bile acid biosynthesis	−0.47	−0.05	−0.22	−0.11	−0.06	−0.03
Primary immunodeficiency	0.09	−0.07	−0.23	−0.1	0.06	0.18
Progesterone mediated oocyte maturation	0.22	0.26	0.08	0.17	0.12	−0.12
Propanoate metabolism	−0.22	−0.21	−0.16	−0.1	−0.15	−0.11
Proteasome	0.17	0.41	0.56	0.1	0.29	0.23
Pyrimidine metabolism	0.28	0.67	0.22	0.41	0.09	0.04
Regulation of actin cytoskeleton	0.13	0.09	0.17	−0.09	0.17	−0.08
Renin angiotensin system	−0.08	0.04	−0.25	−0.27	−0.2	0.13
Riboflavin metabolism	0.13	0.05	0.2	−0.04	0.05	0.19
Rna degradation	0.35	0.21	0.38	0.15	0.18	0.05
Rna polymerase	0.21	0.55	0.07	0.25	0.1	0.09
Small cell lung cancer	0.13	0.13	0.15	−0.08	0.18	−0.07
Spliceosome	0.22	0.26	0.21	0.17	0.09	−0.07
Steroid hormone biosynthesis	−0.19	−0.29	−0.08	−0.27	0.02	−0.13
Thyroid cancer	0.1	0.02	0.2	0.23	0.11	−0.14
Tryptophan metabolism	−0.29	−0.19	−0.18	−0.22	−0.09	0.01
Type II diabetes mellitus	−0.16	−0.03	−0.11	−0.1	0.12	−0.18
Type I diabetes mellitus	−0.09	−0.41	−0.29	−0.23	−0.09	0.07
Tyrosine metabolism	−0.3	−0.27	−0.42	−0.12	−0.15	−0.08
Ubiquitin mediated proteolysis	−0.08	0.18	0.14	0.27	0.14	−0.12
Valine leucine and isoleucine degradation	−0.21	−0.19	−0.01	−0.16	−0.14	−0.12
Vasopressin regulated water reabsorption	0.14	0.14	0.09	0.05	−0.05	−0.18
Viral myocarditis	0.11	−0.23	−0.15	−0.18	0.01	0.11
Wnt signaling pathway	0.05	0.04	0.17	0.27	0.05	−0.02

Pathway	Correlation coefficient				
	APITD1	CENPT	CENPU	CENPW	STRA13
Alanine aspartate and glutamate metabolism	0.05	−0.13	0.01	0.09	0.04
Aldosterone regulated sodium reabsorption	0.1	−0.35	−0.05	−0.23	−0.11
Aminoacyl trna biosynthesis	0.02	0.1	−0.12	−0.13	0
Antigen processing and presentation	0.23	−0.25	−0.18	0.17	0.2
Arachidonic acid metabolism	0.04	−0.15	−0.19	−0.29	−0.15
Arrhythmogenic right ventricular cardiomyopathy arvc	0.02	−0.1	0.14	0.03	0.02
Ascorbate and aldarate metabolism	0.15	−0.18	−0.04	−0.04	−0.3
Axon guidance	0.15	0.02	0.06	0	−0.02
Basal cell carcinoma	−0.06	0.28	0.11	−0.14	−0.06
Basal transcription factors	−0.09	−0.02	0.05	0.23	0.14
Base excision repair	0.01	0.12	0.03	0.03	0.05
Butanoate metabolism	0.05	−0.21	−0.12	−0.21	−0.12
Cell adhesion molecules cams	0.02	−0.15	−0.11	−0.09	0.07
Cell cycle	0.04	0.09	0.19	0.28	0.19
Chronic myeloid leukemia	0.16	0.02	0.03	0.01	0.08
Colorectal cancer	0.12	0.03	0.17	0.01	0.03
Dna replication	0.37	0.05	0.03	0.07	0.03
Dorso ventral axis formation	0.1	−0.18	0.15	−0.02	−0.12
Drug metabolism cytochrome p450	−0.12	−0.13	−0.49	−0.15	−0.24
Fatty acid metabolism	0.06	0	0.03	−0.21	−0.13
Gap junction	0.1	−0.12	0.12	0.12	−0.12
Glutathione metabolism	0.04	−0.33	0.01	−0.34	−0.33
Glycosphingolipid biosynthesis ganglio series	0.12	−0.2	0.1	−0.13	0.09
Glycosphingolipid biosynthesis lacto and neolacto series	0.05	−0.1	−0.05	−0.14	−0.12
Gnrh signaling pathway	0.18	−0.11	0.04	−0.04	−0.05
Histidine metabolism	−0.08	−0.32	−0.14	−0.18	−0.12
Homologous recombination	0.13	0.12	0.11	0.16	0.36

(*continued*)

Table 3
(continued).

Pathway	Correlation coefficient				
	APITD1	CENPT	CENPU	CENPW	STRA13
Inositol phosphate metabolism	0.05	−0.13	0.25	0.32	0.13
Intestinal immune network for iga production	0.05	−0.14	−0.35	0.23	0.22
Leukocyte transendothelial migration	0.09	−0.1	0.01	−0.1	−0.06
Linoleic acid metabolism	0.03	−0.18	−0.12	−0.09	−0.26
Long term depression	0.02	−0.3	0.02	−0.03	0.09
Long term potentiation	0.13	0.02	0.02	−0.02	0.06
Lysine degradation	−0.1	−0.11	−0.06	−0.1	−0.1
Melanogenesis	0.08	−0.12	0.09	−0.04	0.01
Melanoma	−0.01	−0.11	−0.03	−0.14	−0.04
Metabolism of xenobiotics by cytochrome p450	−0.11	−0.14	−0.47	−0.11	−0.2
Mismatch repair	−0.06	−0.02	0.04	−0.03	0.16
Neuroactive ligand receptor interaction	0.03	−0.39	−0.01	−0.05	−0.16
Neurotrophin signaling pathway	0.02	−0.02	0.02	−0.03	−0.09
Non homologous end joining	0.1	0.19	0.19	0.17	0.23
Nucleotide excision repair	0.04	0.06	0.03	0.1	0.17
Olfactory transduction	−0.11	−0.07	−0.16	0.02	−0.14
One carbon pool by folate	−0.25	0.1	0.16	0.28	0.26
Oocyte meiosis	−0.04	−0.15	0.15	0.18	0.1
P53 signaling pathway	−0.05	−0.04	0.16	0.3	0.15
Pancreatic cancer	0.08	−0.03	0.12	0.12	0.03
Pantothenate and coa biosynthesis	0.23	−0.12	0.16	0.07	0.11
Pathogenic escherichia coli infection	0.16	−0.01	0.13	0.12	0.13
Pathways in cancer	0.1	−0.12	0.1	0	0.02
Primary bile acid biosynthesis	0.03	−0.11	−0.21	−0.41	−0.15
Primary immunodeficiency	0.23	−0.13	0.02	0.16	0.18
Progesterone mediated oocyte maturation	−0.01	−0.13	0.2	0.08	0.16
Propanoate metabolism	0.25	−0.05	−0.16	−0.21	−0.2
Proteasome	0.23	−0.22	0.23	0.37	0.12
Pyrimidine metabolism	0.14	−0.19	0.02	0.19	0.05
Regulation of actin cytoskeleton	0	−0.12	0.13	0.11	0.06
Renin angiotensin system	0.2	−0.18	0.03	0.05	−0.23
Riboflavin metabolism	0.22	0.01	−0.06	0.22	0.21
Rna degradation	0.03	0.1	0.09	0.39	0.13
Rna polymerase	0.12	0.06	0.07	0.05	0.08
Small cell lung cancer	0.12	−0.12	0.1	0	0.06
Spliceosome	−0.05	0.08	0.04	0.09	0.12
Steroid hormone biosynthesis	0.1	−0.12	−0.06	−0.05	−0.24
Thyroid cancer	0.09	−0.15	0.14	−0.1	0.17
Tryptophan metabolism	0.09	−0.11	−0.21	−0.4	−0.08
Type II diabetes mellitus	0.02	−0.28	0.07	−0.04	−0.12
Type I diabetes mellitus	0.01	−0.26	−0.21	0.11	0.13
Tyrosine metabolism	−0.25	−0.1	−0.26	−0.46	−0.17
Ubiquitin mediated proteolysis	0.08	0.03	0.09	−0.1	0.23
Valine leucine and isoleucine degradation	0.22	−0.35	−0.02	−0.15	−0.13
Vasopressin regulated water reabsorption	0.05	−0.2	0.22	0.31	0.09
Viral myocarditis	0.13	−0.25	0.04	0.09	0.07
Wnt signaling pathway	0.12	0.03	0.22	−0.01	−0.07

APITD1 = apoptosis-inducing TAF 9-like domain 1, CENPA = centromere protein A, CENPC = centromere protein C, CENPH = centromere protein H, CENPI = centromere protein I, CENPK = centromere protein K, CENPL = centromere protein L, CENPM = centromere protein M, CENPN = centromere protein N, CENPO = centromere protein O, CENPP = centromere protein P, CENPQ = centromere protein Q, CENPT = centromere protein T, CENPU = centromere protein U, CENPW = centromere protein W, ITGB3BP = integrin subunit beta 3 binding protein, STRA13 = stimulated by retinoic acid 13.

APITD1, CENPT, CENPU, and STRA13 were most negative related to the stromal cell. In the same measure, CENPW, CENPH, CENPA, CENPU, and CENPM were most positive related to the immune cell, and ITGB3BP, APITD1, CENPP, CENPQ, CENPT, and CENPC were most negative related to the immune cell (A and B).

Next, the connection between CCAN expression and immune cells infiltration was further be explored. We purposely selected the three genes with the ultimate positive correlation score of immune cells in high expression CCAN and the three genes with the ultimate negative correlation score of immune cells in low expression CCAN as the research objects. The results showed that high expression CCAN mainly related to macrophages M0, macrophages M1, mast cells activated, mast cells resting, monocytes, neutrophils, T cells regulatory in pan-cancer, besides, low expression CCAN mainly related to NK cells resting, T cells follicular helper, T cells CD8, neutrophils, macrophages M0, T cells CD4 memory activated (C and D).

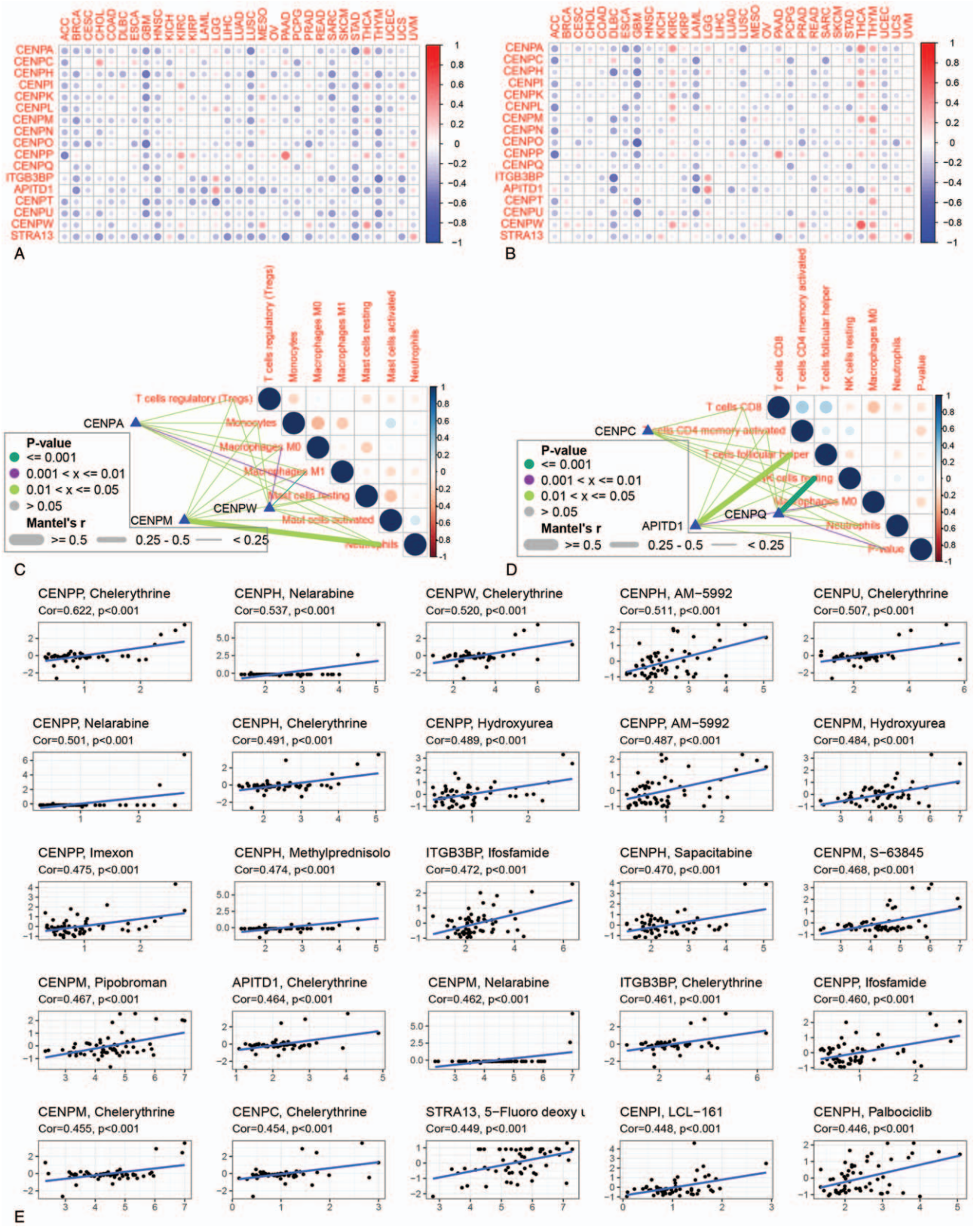

Figure 5. The immune and drug correlation with CCAN. (A and B) Correlation between CCAN expression and immune microenvironment of stromal cell (A) or immune cell (B). (C and D) Correlation between CCAN expression and immune cells infiltration of immune positive related genes (C) or immune negative related genes (D). (E) Correlation between CCAN expression and drug sensitivity.

Using the cellminer database, we inspected the association between CCAN expression and drug sensitivity. There are 632 drugs associated with these 17 CCAN genes, of which the highest correlation coefficient combinations are listed in Figure 5E, including chelerythrine, nelarabine, hydroxyurea, and so on.

4. Discussion

At metaphase of mitosis, the chromosomes oscillate regularly and harmoniously to the metaphase plate to form the equator. The kinetochore is located on both sides of the centromere. By connecting with the k-fiber, it generated the plus-end directed microtubule-motor forces to draw the chromosome to the spindle on both sides. Protein complex CCAN is the core conserved part of centromere and engaged a crucial role in this system. On the one hand, CCAN is composed of a variety of proteins, which are connected with each other and serve as a scaffold between the microtubule plus ends and the centromeric DNA. On the other hand, CCAN, as the core regulator of kinetochore-microtubule dynamics, can specifically identifies spreading microtubules, reduced the transition of kinetochore microtubule and terminal, that is, prevent the incontrollable and speedy dynamic fluctuation of kinetochore microtubule.[13] In our study, we revealed the panoramic picture of CCAN in pan-cancer and established a new understanding of CCAN.

First of all, we explored the different expression profile of CCAN in pan-cancer from mRNA, protein and cell level. The UCSC xena results were divided into three groups by the mRNA expression. There were CENPW, CENPH, CENPA, CENPU, and CENPM in the high expression group, CENPL, CENPI, CENPO, CENPK, and CENPN, STRA13 in the medium expression group and APITD1, CENPP, ITGB3BP, CENPQ, CENPT, and CENPC in the low expression group. Without divergence, the difference of oncomine expression was implicit, but the trend of differential expression of CCAN was consistent with our finding in UCSC xena. In the protein level, the trend of expression was quite lower than that of mRNA, it may be because the sample size is small, or the amount of protein will change dynamically according to the needs of mitosis, which needs further study. In the cell level, we analyzed the cell expression profile of CCAN in breast, colorectal, endometrial, liver, lung, and prostate. The results were basically consistent with the mRNA expression in tissues. Previous clinical studies have shown that CENPA, CENPK, CENPL, CENPM, CENPN, CENPU, and CENPW was consistently overexpressed in many human cancers, which proved our results.[14]

Cox analysis showed that some CCAN were associated with descending survival. Therefore, we try to speculate on the potential causes of CCAN in the aspects of signaling pathway, gene variation and immune environment. By analyzing the interrelationship between CCAN expression and kegg pathway, it was concluded that CCAN was involved in 184 enrichment pathways, 74 of which are common pathways. According to the kegg classification, the 74 pathways involved are cellular processes (10.6%), environmental information processing (10.1%), genetic information processing (17.6%), human diseases (18.9%), metabolism (32.4%), and organismal systems (20.3%). These results indicated that CCAN complexes not only cooperated with each other to complete signal transduction, but also played their own roles independently. According to the early literature, CENPN could promote oral cancer by strengthening cell cycle. On the contrary, G1 phase stagnated when CENPN decreased, accompanied with augmentation of p21 CIP1, p27 kip1 and abatement of cyclin D1, CDK2, CDK4.[15] In addition, CENPA promoted chromosome instability and tumorigenesis by phosphorylation on serine 18.[16] Further research confirmed that CENPK promote hepatocellular carcinoma with very poor prognosis through activated the AKT and MDM2 protein into phosphorylation in the position of tyrosine, but the TP53 protein tyrosine out of phosphorylation.[17] In general, the results mentioned above indicated that some CCAN directly participated in cancer signaling pathway, which leaded to poor prognosis.

It is all to know that loss of CCAN integrity can lead to centromere dysfunction, resulting in relocation of microtubules minus-ends, dispersion of pericentriolar material during mitosis, and chromosome separation errors.[18] However, in the study of mutation frequency and copy number variation frequency, CCAN family genes are relatively conservative and stable. Only CENPI and CENPC mutations are obvious in uterine corpus endometrial carcinoma, and copy number changes are obvious in kidney chromophobe and uterine carcinosarcoma. This showed the importance of CCAN in epigenetics from another point of view. It is inferred that the variation of CCAN gene is not the main cause but the important cause of cancer.

At present, the research on CCAN immunity mainly focuses on autoimmune systemic sclerosis, however, the reports about CCAN in the field of tumor immunity were deficiency. In this study, we analyzed the association between CCAN expression and tumor microenvironment, the results showed significant correlation in stromal cell and immune cell, which indicated the potential metastasis capability of tumor. Furthermore, in order to identify specific immune cells that regulate tumor microenvironment, we selected the high expression group and low expression group respectively for immunocyte infiltration analysis, the results show that the two groups of genes are associated with immune cells, but the types of immune cells are mostly different. Previous studies have shown that $CD4^+$ T cell were regulated by CENPC and CENPB were in a resting state before cell division, the content of CENPA was very low at the same time. After entering the cell cycle, T cells were activated and CENPA was increased by reloading.[19] Combined with our progressive studies, we have reason to believe that there is compact linkage across CCAN and tumor immune cell infiltration, and it may be a factor to affect the prognosis of tumor treatment, but the dependence of them is worth further exploration.

The correlation examination between CCAN expression and drug sensitivity in pan carcinoma was listed. Chelerythrine had the highest correlation with multiple genes, which was investigated the effects as a cancer treatment in several studies.[20–22] In terms of signal transduction pathway, chelerythrine was involved with the pathways of Wnt, NF-κB, PI3K/BAD, and so on.[23–25] In terms of tumor immune cell infiltration, chelerythrine was involved with CD4+ and CD8+ T cell infiltration.[26] The above results are highly coincident with our analysis results, which reflects the reliability of the system, and also help to improve the precision treatment of cancer caused by CCAN.

5. Conclusion

To summarize, our study comprehensively analyzed the expression, pathway, immunity and treatment of CCAN in pan-cancer, described the cancer network of CCAN, and revealed the possibility of CCAN as a diagnostic, prognostic biomarker or

therapeutic target in the future. However, there are still many problems to be further explored, such as the asymmetry of protein expression and mRNA expression profile, the correlation with tumor stem cells, hypoxia, and autophagy and so on.

Acknowledgments

This research was supported by: Key laboratory of translational cancer research (Putian university), (Grant No. 2018KF001). We declare that we have no conflict of interest.

Author contributions

Conceptualization: huimei su, Li-he Jiang.
Data curation: huimei su.
Formal analysis: huimei su.
Investigation: huimei su.
Methodology: huimei su.
Project administration: Li-he Jiang.
Resources: Li-he Jiang.
Software: huimei su.
Supervision: Li-he Jiang.
Validation: huimei su.
Visualization: huimei su, Zhuan Wang.
Writing – original draft: huimei su, Yuchun Fan.
Writing – review & editing: Li-he Jiang.

References

[1] Vargiu G, Makarov AA, Allan J, Fukagawa T, Booth DG, Earnshaw WC. Stepwise unfolding supports a subunit model for vertebrate kinetochores. Proc Natl Acad Sci U S A 2017;114:3133–8.
[2] Yan K, Yang J, Zhang Z, et al. Structure of the inner kinetochore CCAN complex assembled onto a centromeric nucleosome. Nature 2019;574: 278–82.
[3] Negrini S, Gorgoulis VG, Halazonetis TD. Genomic instability—an evolving hallmark of cancer. Nat Rev Mol Cell Biol 2010;11:220–8.
[4] Zhang Y, Yang L, Shi J, Lu Y, Chen X, Yang Z. The oncogenic role of CENPA in hepatocellular carcinoma development: evidence from bioinformatic analysis. Biomed Res Int 2020;2020:3040839.
[5] Lu G, Hou H, Lu X, et al. CENP-H regulates the cell growth of human hepatocellular carcinoma cells through the mitochondrial apoptotic pathway. Oncol Rep 2017;37:3484–92.
[6] Wang Q, Yu X, Zheng Z, Chen F, Yang N, Zhou Y. Centromere protein N may be a novel malignant prognostic biomarker for hepatocellular carcinoma. PeerJ 2021;9:e11342.
[7] Rahman MR, Islam T, Gov E, et al. Identification of prognostic biomarker signatures and candidate drugs in colorectal cancer: insights from systems biology analysis. Medicina (Kaunas) 2019;55: doi:10.3390/medicina 55010020.
[8] Zhang S, Xie Y, Tian T, et al. High expression levels of centromere protein A plus upregulation of the phosphatidylinositol 3-kinase/Akt/ mammalian target of rapamycin signaling pathway affect chemotherapy response and prognosis in patients with breast cancer. Oncol Lett 2021; 21:410.
[9] Zhai X, Yang Z, Liu X, Dong Z, Zhou D. Identification of NUF2 and FAM83D as potential biomarkers in triple-negative breast cancer. PeerJ 2020;8:e9975.
[10] He P, Hu P, Yang C, He X, Shao M, Lin Y. Reduced expression of CENP-E contributes to the development of hepatocellular carcinoma and is associated with adverse clinical features. Biomed Pharmacother 2020; 123:109795.
[11] Okumura K, Kagawa N, Saito M, et al. CENP-R acts bilaterally as a tumor suppressor and as an oncogene in the two-stage skin carcinogenesis model. Cancer Sci 2017;108:2142–8.
[12] Hoadley KA, Yau C, Hinoue T, et al. Cell-of-origin patterns dominate the molecular classification of 10,000 tumors from 33 types of cancer. Cell 2018;173:291–304.
[13] McAinsh AD, Meraldi P. The CCAN complex: linking centromere specification to control of kinetochore-microtubule dynamics. Semin Cell Dev Biol 2011;22:946–52.
[14] Zhang W, Mao JH, Zhu W, et al. Centromere and kinetochore gene misexpression predicts cancer patient survival and response to radiotherapy and chemotherapy. Nat Commun 2016;7:12619.
[15] Oka N, Kasamatsu A, Endo-Sakamoto Y, et al. Centromere protein N participates in cellular proliferation of human oral cancer by cell-cycle enhancement. J Cancer 2019;10:3728–34.
[16] Zhang W, Karpen GH, Zhang Q. Exploring the role of CENP-A Ser18 phosphorylation in CIN and tumorigenesis. Cell Cycle 2017;16:2323–5.
[17] Overexpression of centromere protein K (CENP-K) gene in hepatocellular carcinoma promote cell proliferation.pdf.
[18] Gemble S, Simon A, Pennetier C, et al. Centromere dysfunction compromises mitotic spindle pole integrity. Curr Biol 2019;29:3072–80.
[19] Hoffmann S, Izquierdo HM, Gamba R, et al. A genetic memory initiates the epigenetic loop necessary to preserve centromere position. EMBO J 2020;39:e105505.
[20] Yang R, Piperdi S, Gorlick R. Activation of the RAF/mitogen-activated protein/extracellular signal-regulated kinase kinase/extracellular signal-regulated kinase pathway mediates apoptosis induced by chelerythrine in osteosarcoma. Clin Cancer Res 2008;14:6396–404.
[21] Wan KF, Chan SL, Sukumaran SK, Lee MC, Yu VC. Chelerythrine induces apoptosis through a Bax/Bak-independent mitochondrial mechanism. J Biol Chem 2008;283:8423–33.
[22] In vitro and in vivo activity of protein kinase C inhibitor chelerythrine chloride induces tumor cell toxicity and growth delay in vivo.pdf.
[23] Heng WS, Cheah SC. Chelerythrine chloride downregulates beta-catenin and inhibits stem cell properties of non-small cell lung carcinoma. Molecules 2020;25: doi:10.3390/molecules25010224.
[24] Fan L, Fan Y, Liu L, et al. Chelerythrine attenuates the inflammation of lipopolysaccharide-induced acute lung inflammation through NF-kappaB signaling pathway mediated by Nrf2. Front Pharmacol 2018;9:1047.
[25] Yang T, Xu R, Su Q, et al. Chelerythrine hydrochloride inhibits proliferation and induces mitochondrial apoptosis in cervical cancer cells via PI3K/BAD signaling pathway. Toxicol In Vitro 2020;68:104965.
[26] Zhang Q, Tian Y, Duan J, et al. Chelerythrine ameliorates acute cardiac allograft rejection in mice. Transpl Immunol 2016;38:78–83.