Pan-Cancer Analysis Identifies CCDC51 as a Potential Biomarker for Liver Hepatocellular Carcinoma

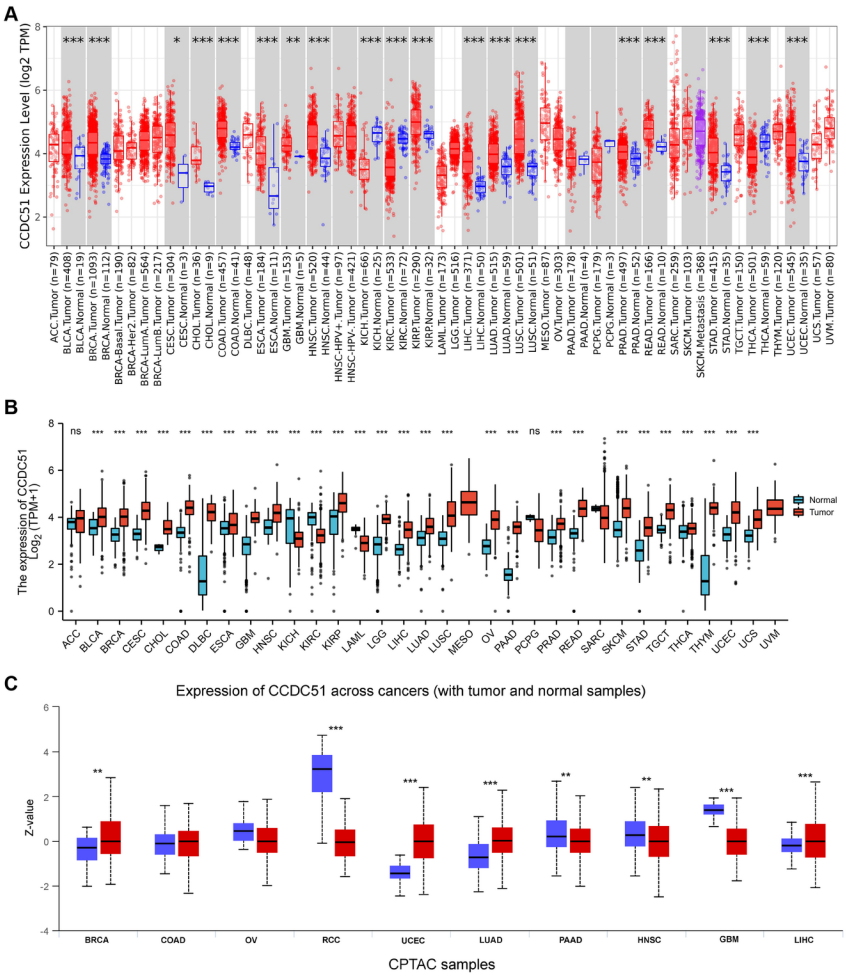
Shaoyi Wang[1], peng chen[2], zhang chen[3], and Xiaoyan Huang[1]

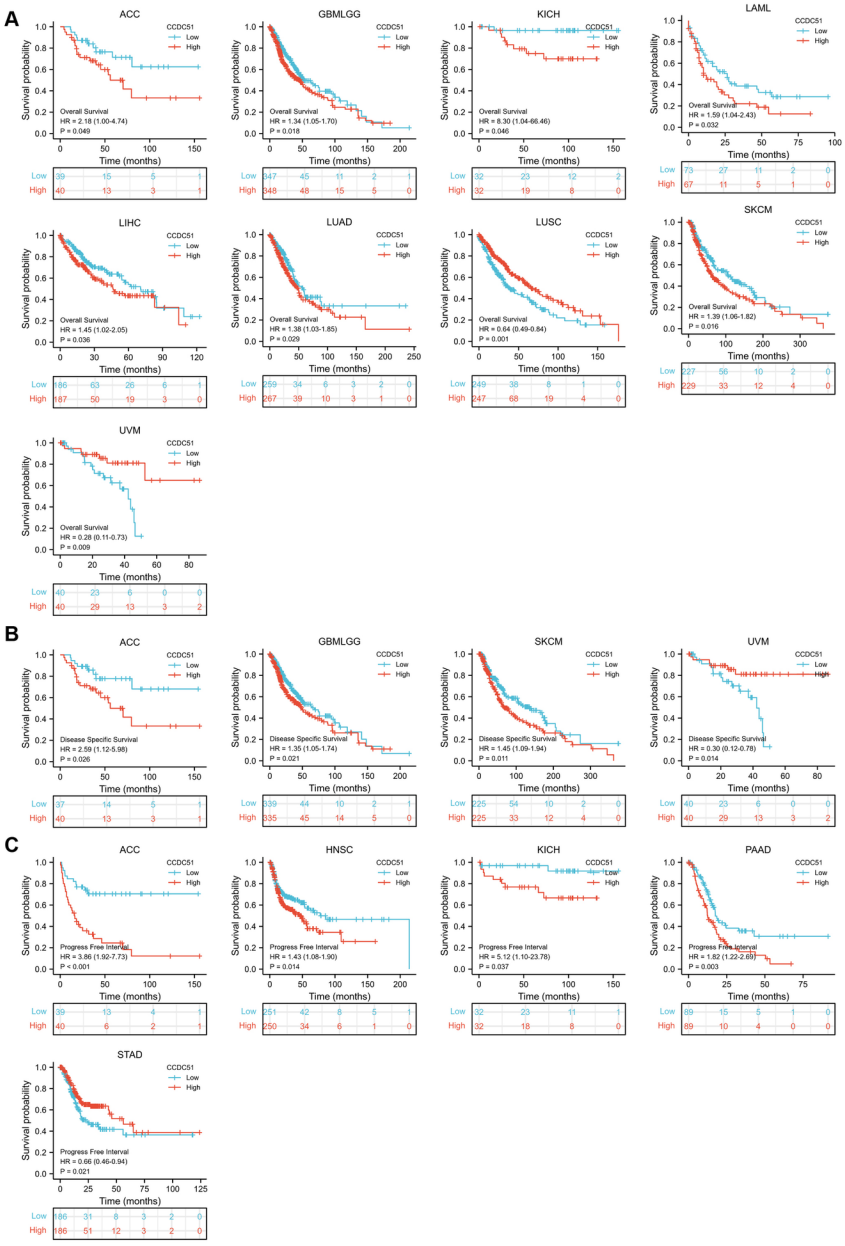
[1]The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University
[2]YONGJIA PEOPLE'S HOSPITAL
[3]The Second Affiliated Hospital, Yuying Children's Hospital of Wenzhou Medical University

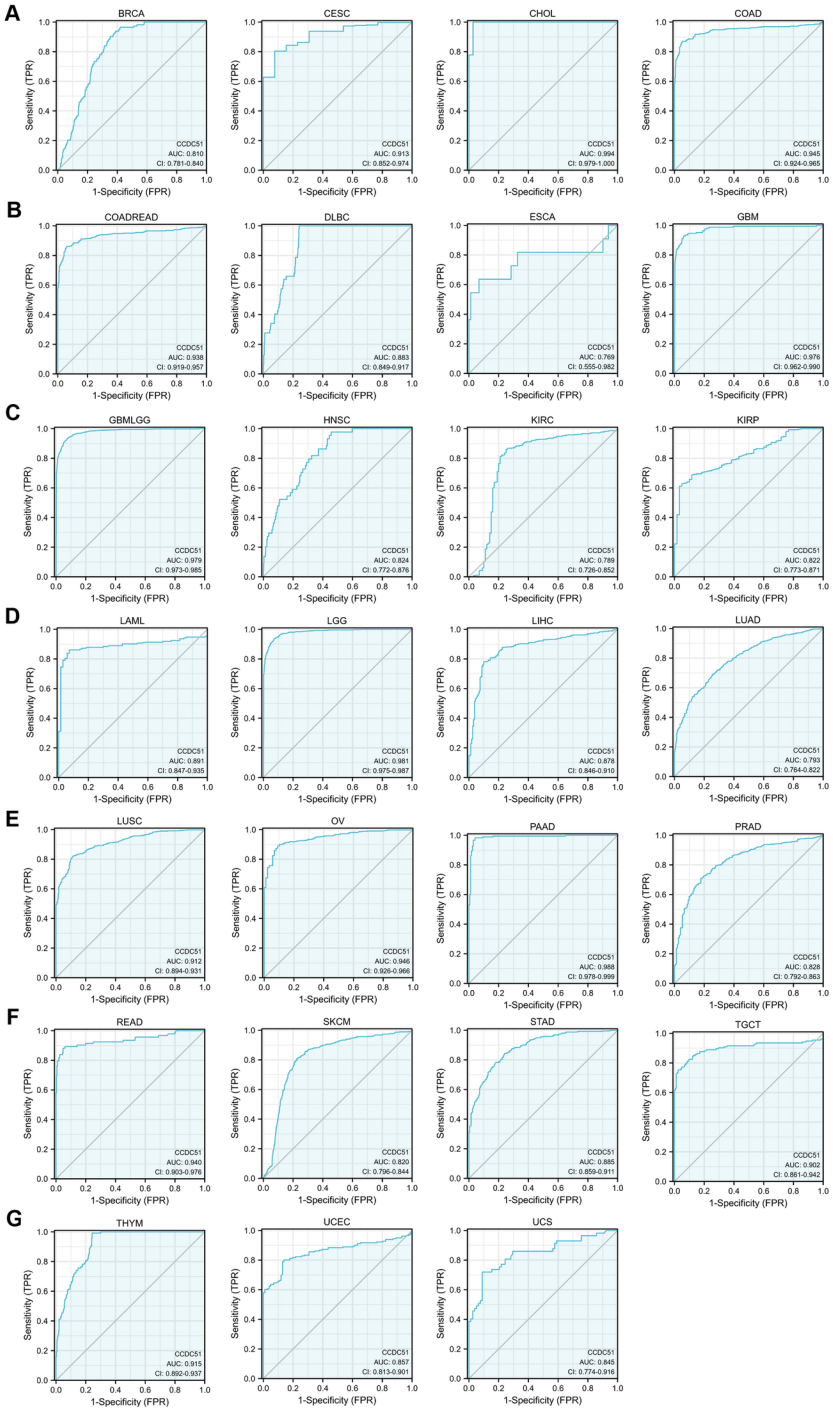
May 30, 2023

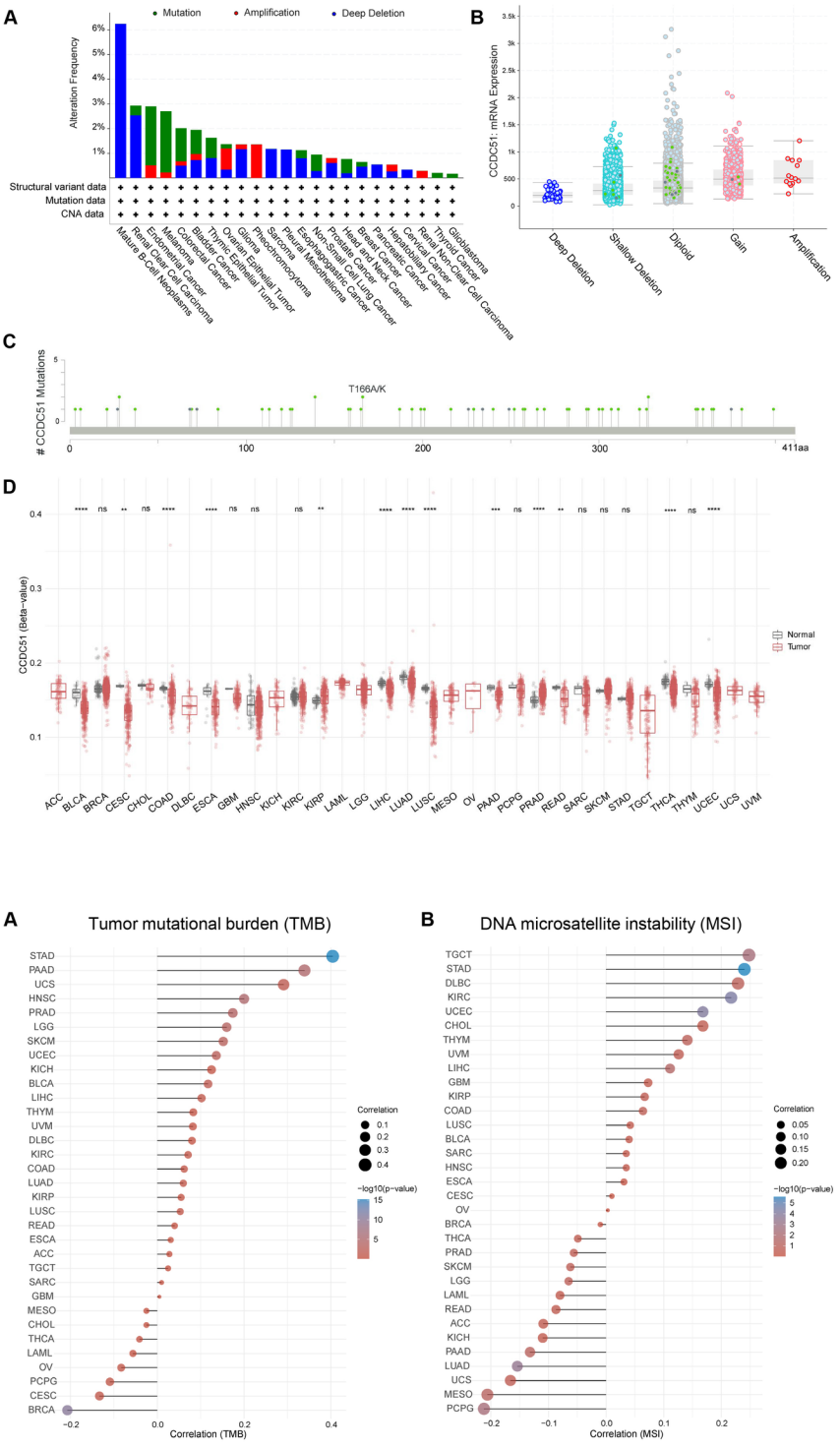
Abstract

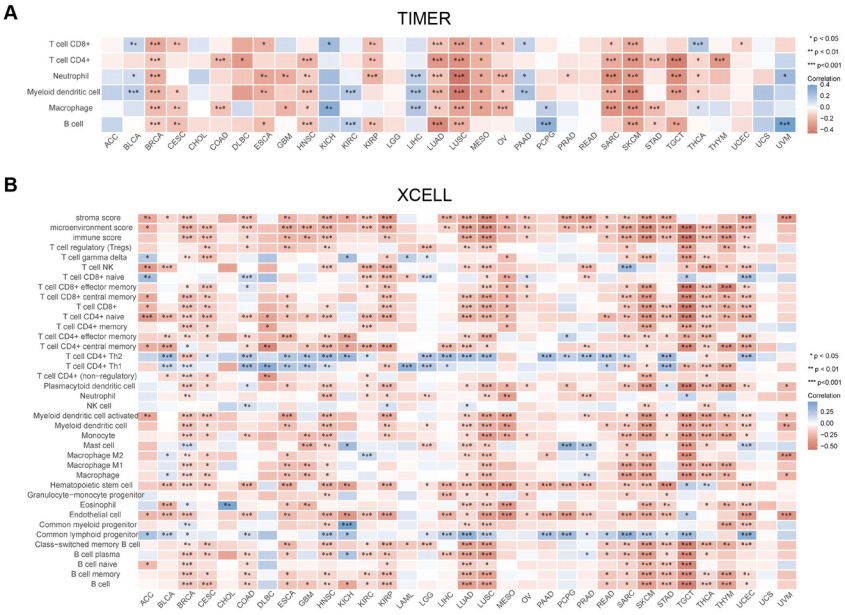
Background: Coiled coil domain containing protein 51 (CCDC51), as the two transmembrane helical domains protein, little is known about the function of it in human cancer. Methods: TCGA, GEPIA, cBioPortal, GTEx, and TIMER were employed to analyze expression, immune cells infiltration, prognostic value, genetic alteration, cancer patients' mutation burden, stem cell stability, and microsatellite instability of CCDC51 . Results: Decreased CCDC51 expression significantly related with poor overall survival (OS) of acute myeloid leukemia (LAML), adrenocortical carcinoma (ACC), glioma (GBMLGG), kidney chromophobe (KICH), liver hepatocellular carcinoma (LIHC), lung squamous cell carcinoma (LUSC), skin cutaneous melanoma (SKCM) and uveal melanoma (UVM), lung adenocarcinoma (LUAD), disease-specific survival (DSS) of ACC, GBMLGG, SKCM and UVM, and progression-free interval (PFI) of ACC, KICH and pancreatic adenocarcinoma, squamous cell carcinoma of the head and neck (HNSC). In human cancer, immune cell infiltration, and tumor microenvironment, CCDC51 expression is associated with MSI, RNA modifications, and diverse cancer drug sensitivity. There is potential for it to be an independent factor contributing to OS in LIHC. CCDC51 was an independent factor for LIHC prognosis in Cox regression and nomogram analysis. The results of the the Kyoto Encyclopedia of Genes and Gene Ontology and Genomes indicated that CCDC51 was involved in aminoacyl-tRNA biosynthesis, RNA transport, colorectal cancer, Wnt signaling pathway, mismatch repair, mitochondrion, pseudo uridine synthase activity, mRNA processing, mitochondrial matrix, regulation of mRNA stability, and mitochondrial inner membrane. Conclusion: Our research can provide new-insights for LIHC prognostic biomarkers of CCDC51.

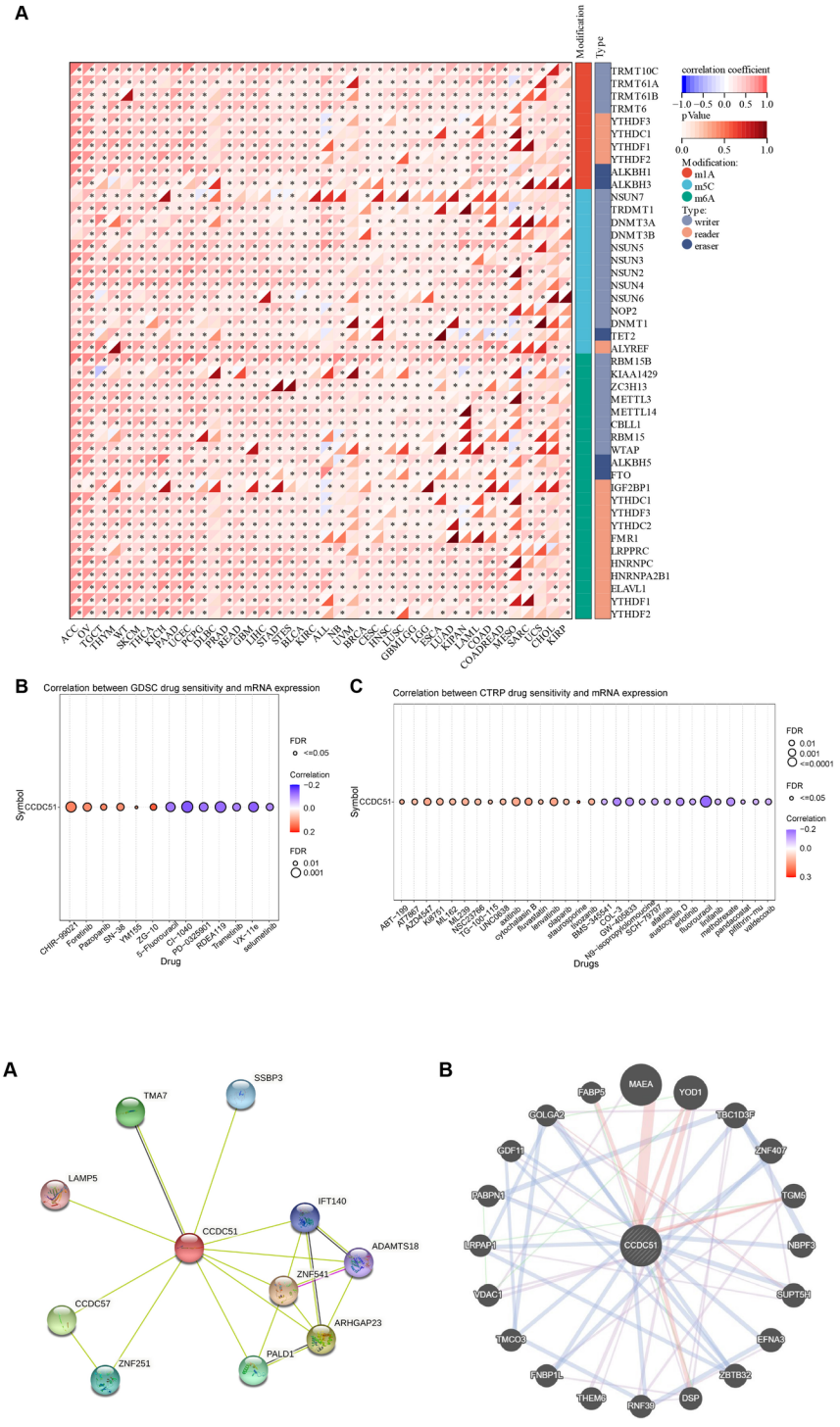


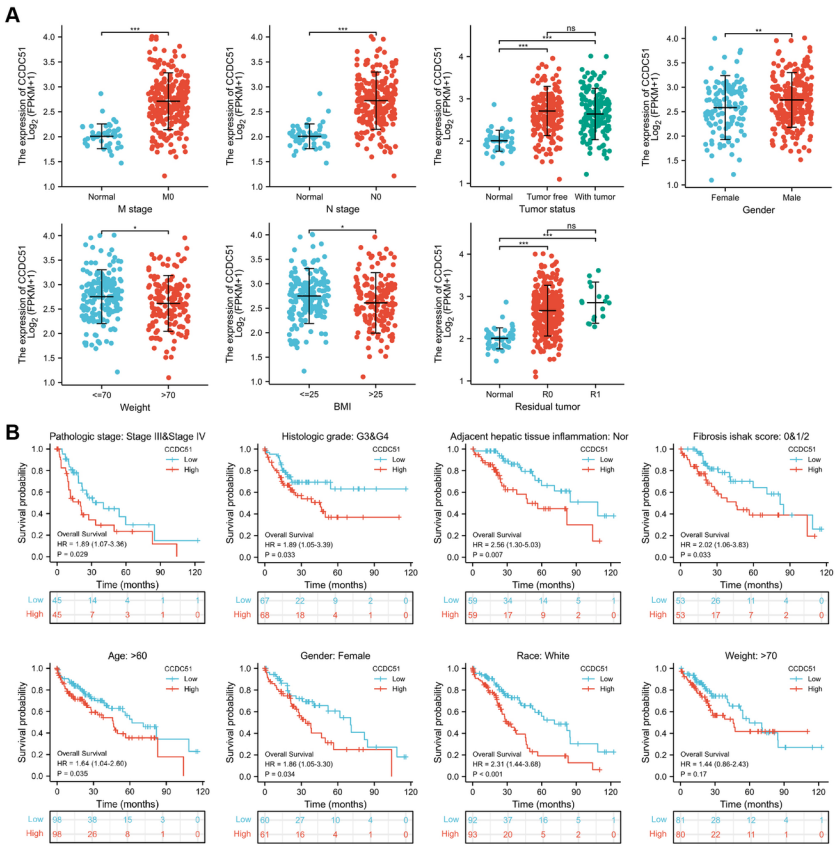


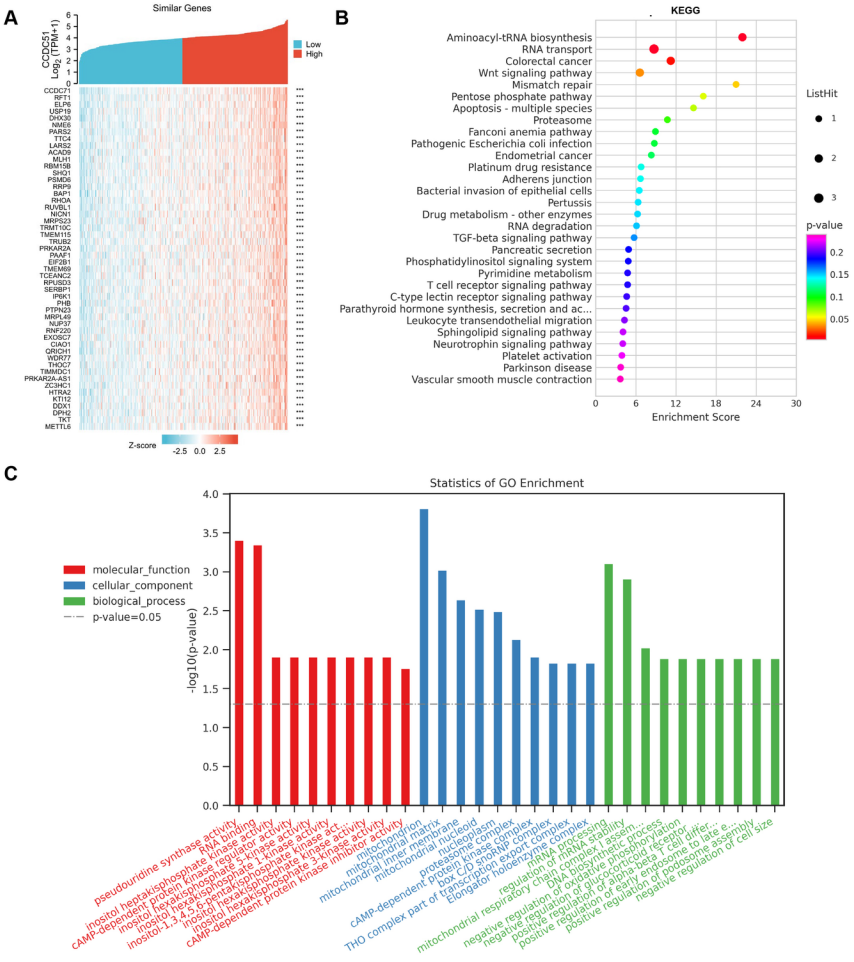












Pan-Cancer Analysis Identifies CCDC51 as a Potential Biomarker for Liver Hepatocellular Carcinoma

Shaoyi Wang[1], Peng Chen[2], Zhang Chen[1], Xiaoyan Huang [1*]

[1] The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, NO.109, Xueyuan West Road, Wenzhou, Zhejiang, China

[2] Yongjia People's Hospital, NO.283, Yongjian Road, Yongjia County, Wenzhou, Zhejiang, China

*** Correspondence:**

Xiaoyan Huang

The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, NO.109, Xueyuan West Road, Wenzhou, Zhejiang, China

Email: **13868396252@163.com; 221180@wzhealth.com**

Abstract

Background: Coiled coil domain containing protein 51 (CCDC51), as the two transmembrane helical domains protein, little is known about the function of it in human cancer.

Methods: TCGA, GEPIA, cBioPortal, GTEx, and TIMER were employed to analyze expression, immune cells infiltration, prognostic value, genetic alteration, cancer patients' mutation burden, stem cell stability, and microsatellite instability of CCDC51 .

Results: Decreased CCDC51 expression significantly related with poor overall survival (OS) of acute myeloid leukemia (LAML), adrenocortical carcinoma (ACC), glioma (GBMLGG), kidney chromophobe (KICH), liver hepatocellular carcinoma (LIHC), lung squamous cell carcinoma (LUSC), skin cutaneous melanoma (SKCM) and uveal melanoma (UVM), lung adenocarcinoma (LUAD), disease-specific survival (DSS) of ACC, GBMLGG, SKCM and UVM, and progression-free interval (PFI) of ACC, KICH and pancreatic adenocarcinoma, squamous cell carcinoma of the head and neck (HNSC). In human cancer, immune cell infiltration, and tumor microenvironment, CCDC51 expression is associated with MSI, RNA modifications, and diverse cancer drug sensitivity. There is potential for it to be an independent factor contributing to OS in LIHC. CCDC51 was an independent factor for LIHC prognosis in Cox regression and nomogram analysis. The results of the the Kyoto Encyclopedia of Genes and Gene Ontology and Genomes indicated that CCDC51 was involved in aminoacyl-tRNA biosynthesis, RNA transport, colorectal cancer, Wnt signaling pathway, mismatch repair, mitochondrion, pseudo uridine synthase activity, mRNA processing, mitochondrial matrix, regulation of mRNA stability, and mitochondrial inner membrane.

Conclusion: Our research can provide new-insights for LIHC prognostic biomarkers of CCDC51.

Keywords: Pan-cancer, CCDC51, Liver hepatocellular carcinoma, Prognostic value, RNA modification, Immune cell infiltration

Background

Liver hepatocellular carcinoma (LIHC) is a multiple malignant tumor in the world[1]. The mortality rate of malignant tumor ranks fourth in the world, and ranks second in China[1,2]. There has been a gradual upward trend last several years. The onset of liver cancer is hidden, most of the patients often miss the best time for surgical treatment[3]. Relying solely on chemotherapy, whether single drug or combined radiotherapy and chemotherapy, the effect is not ideal, resulting in poor efficacy and high mortality of LIHC[4]. Elucidating the specific biomarkers in the process of LIHC development has critical research significance for early diagnosis and individual drug use of LIHC patients.

Coiled coil domain containing protein 51 (CCDC51), presents the pore-forming subunit of a mitoK (ATP) channel, alias MITOK [5]. It has been reported CCDC51 related to retinal disease [5]. However, the CCDC51 role in pan-cancer diagnosis, prognosis, RNA modification, and immune regulation remains unclear.

In this study, we investigated the role of CCDC51 in pan-cancer, and we clarified the relationship between the expression of CCDC51 and the expression of RNA modification regulators, DNA microsatellite instability (MSI), immune cell infiltration in human cancer, and drug sensitivity, tumor mutational burden (TMB). It is a promising molecular target for LIHC as well as a potential biomarker for diagnosis and prognosis of different cancer types .

Materials and methods

Pan-cancer expression of CCDC51

Tumor Immune Estimation Resource (TIMER) (https://cistrome.shinyapps.io/timer/)[6], Cancer Genome Atlas (TCGA) (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga) , Genotype-Tissue Expression (GTEx) database and UALCAN database (http://ualcan.path.uab.edu/)[7] were used to examine the expression of CCDC51 in pan-cancer tissue (ns, $p \geq 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

The prognosis and clinical information of CCDC51 in pan-cancer

We employed the GEPIA databases (http://gepia.cancer-pku.cn/)[8] and PrognoScan databases (http://dna00.bio.kyutech.ac.jp/PrognoScan/index.html)[9] to examine the OS, DSS and PFI of CCDC51 in pan-cancer. (ns, $p \geq 0.05$; *, $p < 0.05$; **, $p < 0.01$; ***, $p < 0.001$).

Mutations of CCDC51 in pan-cancer

We conducted on the pan-cancer CCDC51 gene mutation information via cBioPortal (https://www.cbioportal.org/)[10].

The relationship between RNA modification of CCDC51 in pan-cancer

From the UCSC databases (https://xenabrowser.net/), we downloaded the TCGA TARGET GTEx (PANCAN,N=19131,G=60499) pan-cancer data set. From the previous study, we extracted the expression data of ENSG00000164051 (CCDC51) gene in each sample [11].

GeneMANIA and STRING databases

In order to construct the gene-gene and protein-protein interactions network of CCDC51, GeneMANIA (http://www.genemania.org) and STRING (https://string-db.org/) were used.

Correlation between CCDC51 and cancer drug sensitivity

GDSC (www.cancerRxgene.org) and CTRP (http://portals.broadinstitute.org/ctrp/) databases were used to analyze the correlation between CCDC51 and drug susceptibility [12,13].

Results

CCDC51 was differentially expressed in pan-cancer

First, TIMER database analysis was excavated to examine the CCDC51 expression level in pan-cancers. CCDC51 expression was increased in breast invasive carcinoma (BRCA) and cholangiocarcinoma (CHOL), cancer of the cervical squamous cell and endocervical adenocarcinoma (CESC), bladder urothelial carcinoma (BLCA), glioblastoma multiforme (GBM), head and neck squamous cell carcinoma (HNSC), colon adenocarcinoma (COAD), liver hepatocellular carcinoma (LIHC), esophageal

carcinoma (ESCA), kidney renal papillary cell carcinoma (KIRP), prostate adenocarcinoma (PRAD), lung adenocarcinoma (LUAD), lung squamous cell carcinoma (LUSC), rectum adenocarcinoma (READ), uterine corpus endometrial carcinoma (UCEC) and stomach adenocarcinoma (STAD) tissues compared with adjacent normal tissues. Furthermore, CCDC51 expression was low in kidney renal clear cell carcinoma (KIRC), kidney chromophobe (KICH), and thyroid carcinoma.(Figure 1A).

Next, we combined the GTEx and TCGA database, and the results proved that CCDC51 expression was significantly increased in COAD, BRCA, CESC, BLCA, CHOL, ESCA, lymphoid neoplasm diffuse large B-cell lymphoma (DLBC), KIRP, GBM, pancreatic adenocarcinoma (PAAD), HNSC, LUAD, brain lower grade glioma (LGG), LIHC, READ, STAD, LUSC, ovarian serous cystadenocarcinoma (OV), skin cutaneous melanoma (SKCM), UCEC, esticular germ cell tumors (TGCT), THCA, thymoma (THYM), and uterine carcinosarcoma (UCS) than in paired adjacent normal (Figure 1B).

To investigate the expression leve of CCDC51 protein in human cancers, we used UALCAN database analysis and found that CCDC51 was highly expressed in LIHC, BRCA, UCE and LUAD. In addition, we found that CCDC51 was down-regulated in RCC, PAAD, and GBM (Figure 1C).

Prognosis values of CCDC51 in pan-cancer

The prognostic ability of CCDC51 in pan-cancer was examined due to the fact that it is expressed differently in many cancer types. Increased CCDC51 expression correlated with poor OS in ACC, LIHC, glioma (GBMLGG), KICH, LAML, LUSC, SKCM, LUAD and UVM (Figure 2A), and poor DSS in ACC, GBMLGG, SKCM, UVM (Figure 2B), and poor PFI in ACC, HNSC, KICH, and PAAD (Figure 2C).

CCDC51 could act as a potential biomarker in pan-cancer

A biomarker for pan-cancer using CCDC51 is being investigated. The analysis of ROC curve was conducted, and the results indicated that CCDC51 can be used as a high sensitivity and specificity biomarker (AUC>0.75) for diagnosing BRCA, CESC, CHOL, COAD, colon adenocarcinoma, rectum adenocarcinoma, esophageal

carcinoma (COADREAD), brain lower grade glioma (LGG), DLBC, ESCA, GBMLGG, GBM, HNSC, KIRC, OV, KIRP, LAML, LIHC, LUAD, LUSC, PAAD, TGCT, PRAD, READ, SKCM, STAD, THYM, UCEC and UCS (Figure 3A–E).

Gene mutation landscape and DNA methylation of CCDC51 in pan-cancer

We download the mutational data of CCDC51 from the cBioPortal. In mature B-cell neoplasms, KIRC, endometrial cancer, and melanoma, mutation rates were higher than in other cancers (Figure 4A), and amplification was the most common alteration type. Using the cBioPortal database, 51 missense sites and 7 truncation sites between amino acids 0 and 411 were identified in CCDC51 (Figure 4C). CCDC51 expression was negatively correlated with DNA methylation in BLCA, CESC, COAD, ESCA, KIRP, LIHC, LUAD, LUSC, PAAD, PRAD, READ, THCA and UCEC (Figure 4D). It appears that CCDC51 genetic alterations and DNA methylation affect its prognostic ability.

TMB and MSI analysis of CCDC51 in pan-cancer

TMB has been recognized as a specific and sensitive biomarker for immune checkpoint inhibitor responses by scholars[14-16]. The expression of CCDC51 and TMB of pan-cancer was investigated. In STAD, PAAD, UCS, HNSC, PRAD, LGG, SKCM, UCEC, KICH, BLCA, LIHC, THYM, UVM, DLBC, KIRC, COAD, LUAD, KIRP, LUSC, READ and ESCA, CCDC51 expression was positively correlated with TMB. BRCA, CESC, PCPG, OV, LAML, THCA and CHOL were negatively correlated with TMB (Figure 5A).

Lack of DNA repair activity causes a hypervariable state of DNA sequence[17]. We also analyzed the correlation between CCDC51 expression and MSI in pan-cancer. The expression of CCDC51 was positively correlated with MSI in TGCT, STAD, DLBC, KIRC, UCEC, and CHOL, THYM, UVM, LIHC, GBM, KIRP, COAD, LUSC, BLCA, SARC, HNSC, and ESCA, whereas PCPG, MESO, UCS, LUAD, PAAD, KICH, ACC, READ, LAML, LGG, SKCM, PRAD, and THCA were negatively correlated with MSI (Figure 5B). Collectively, these results indicate that CDCC51 influences antitumor immunity.

Pan-Cancer Immune Cell Infiltration of CCDC51

Cancer progression relies heavily on immune cells [18]. Using the TIMER database, the infiltration levels of T cells CD8+, CD4+, neutrophils, myeloid dendritic cells, macrophages, expression of B cells and CCDC51 were examined in 32 types of cancers. In 27 types of cancer, CCDC51 expression was significantly correlated with the six major immune cells (Figure 6A). In our study of the relationship between CCDC51 expression and immune cell subtypes, we found that CCDC51 expression was significantly correlated with matrix score in 22 cancers, microenvironment score in 24 cancers, immune score in 19 cancers, and immune score in 38 cancers (Figure 6b).The results showed a close relationship between CCDC51 expression and immune cell infiltration in pan-cancer.

RNA modification and drug sensitivity analysis of CCDC51 in pan-cancer

RNA modification has an important role in normal development and tumorigenesis[19,20]. TCGA TARGET GTEx was used to assess the correlation between RNA modification regulators and CCDC51 expression in pan-cancer. There is a correlation between CCDC51 expression and RNA modification regulators in pan-cancer. CCDC51 expression was markedly correlated with m1A, m5C and m6A in many types of cancers (Figure 7A).

The correlation between CCDC51 expression and drug sensitivity was evaluated using cancer cell lines from the Genomics of Drug Sensitivity in Cancer (GDSC) database and the Cancer Therapeutics Response Portal (CTRP) database. According to the GDSC database, CCDC51 expression was positively correlated with CHIR-99021, Foretinib, Pazopanib, SN-38, YM155, and ZG-10, and negatively correlated with 5-Fluorouracil, CI-1040, PD-0325901, RDEA119, Trametinib, VX-11e, and selumetinib. In the CTRP database, The expression of CCDC51 was positively associated with ABT-199, AT7867, AZD4547, Ki8751, ML162, ML239, NSC23766, TG-100-115, UNC0638, Acitinib and Cytochalasin B, Fluvastatin, Lenvatinib, Olaparib and Staurosporine. Negative correlation was observed with BMS-345541, COL-3, GW-405833, N9-isoproploxine, SCH-79797, Afatinib, Austocystinn D, Elotinib, Fluorouracil, Linifanil, Methotrexate, Pandacostat, Pifithrin−mu, and Valdecoxib.

According to the above results, it is expected that CCDC51 will be a promising cancer therapeutic target because it is significantly related to drug sensitivity in cancer cells.

Interaction network of CCDC51 at the gene and protein levels

Using STRING, we conducted CCDC51 protein–protein interaction network of the seed gene. As expected, several nodes (11) and edges (19) were obtained in the PPI network (Figure 8 A). The gene–gene interaction network of CCDC51 is generated through GeneMANIA as presented in Fig. 8B.

Prognostic analysis of CCDC51 in LIHC

An analysis was conducted to uncover the relationship between CCDC51 expression and LIHC pathology. MN stage, tumor status, gender, weight, BMI, and residual tumour were strongly associated with overexpression of CCDC51 (Figure 9A). High CCDC51 expression was associated with poorer OS for LIHC in most clinical and demographic subgroups. They included pathologic stage (III&IV), histologic grade (G3&G4), adjacent tissue inflammation, fibrosis ishak score, age (>60), gender (female), race(white), weight (>70kg) (Figure 9B). To determine whether CCDC51 level could be used as a prognostic factor in LIHC patients, a multivariate Cox regression analysis was performed. The TCGA LIHC cohort showed that decreased CCDC51 expression and pathological stage were independent prognostic factors(Tables 1–3).

LIHC functions of CCDC51

To analysis KEGG and GO enrichment, top 50 similar genes were captured in GEPIA database (Figures 10A). For the KEGG enrichment, these genes mainly covered in aminoacyl-tRNA biosynthesis, RNA transport, colorectal cancer, Wnt signaling pathway, and mismatch repair (Figures 10B). GO enrichment results showed that these genes mainly covered in mitochondrion, pseudo uridine synthase activity, mRNA processing, mitochondrial matrix, regulation of mRNA stability, and mitochondrial inner membrane (Figures 10C). Findings show that CCDC51 plays a critical role in regulating LIHC malignant progression.

Discussion

Pan cancer analysis is critical to compare heterogeneity between different tumors for

identifying novel cancer biomarkers and therapeutic targets[21]. CCDC51 has been demonstrated to be involved in retinal disease [5]. Currently, no studies have tested whether CCDC51 is associated with cancer prognosis.

In this research, we analyzed CCDC51 expression in pan-cancer. CCDC51 expression was significantly higher in BLCA, CHOL, COAD, CESC, DLBC, ESCA, BRCA, GBM, KIRP, LGG, LUAD, HNSC, LUSC, OV, LIHC, SKCM, PAAD, READ, THYM, UCEC, STAD, TGCT, THCA, and UCS than in paired adjacent normal tissues. In addition, our result indicated high CCDC51with poor OS for LIHC, ACC, GBMLGG, LAML, LUAD, LUSC, KICH, SKCM, and UVM, poor DSS in ACC, GBMLGG, SKCM, and UVM, and poor PFI in ACC, HNSC, KICH, and PAAD. ROC curve analysis indicated that CCDC51 could be used as a biomarker for the diagnosis of different types of cancer, with high sensitivity and specificity. Additionally, our team ensured the relationship between CCDC51 and mutation. We found that mature B-Cell neoplasms patients had the highest alteration frequency of CCDC51 (>6%). We also confirmed that alteration frequencies were 2.94%, 2.9%, 2.7%, and 2.02% in KIRC, endometrial cancer, melanoma, and colorectal cancer, respectively. Furthermore, DNA methylation of CCDC51 affects its prognostic ability. TMB and MSI as an immunotherapy biomarker for cancer immune checkpoint inhibitors [15,16]. CCDC51 was correlated with TMB in 31 cancer types and MSI in 30 cancer types. Next, analyzed gene-gene and protein-protein interaction networks, and 10 proteins interacted with CCDC51, namely, SSBP3, TMA7, LAMP5, CCDC75, ZNF251, IFT140, ZNF541, ARHGAP23, ADAMTS18 and PALD1.

LIHC is a multiple malignant tumor in the world, which is a serious threat to human health[1]. CCDC51 expression levels and pathological stage were independent prognostic factors in the TCGA, LIHC cohort.

In order to gain a deeper understanding of CCDC51's role in LIHC, KEGG and GO enrichment suggested that CCDC51 was involved in aminoacyl-tRNA biosynthesis, RNA transport, colorectal cancer, Wnt signaling pathway, mismatch repair, mitochondrion, pseudo uridine synthase activity, mRNA processing, mitochondrial matrix, regulation of mRNA stability, and mitochondrial inner membrane.

CCDC51 and pan-cancer are now better understood, but there are still some unanswered questions. First of all, although we have explored the relationship between CCDC51 and immune infiltration in pan-cancer, there is no experimental evidence to confirm its role in immune regulation of TME in cancer. Secondly, in vitro and in vivo studies of CCDC51's potential role in cancer development are needed.

Conclusion

To sum up, we revealed the potential function of CCDC51 in pan-cancer. Additionally, CCDC51 expression was associated with pan-cancer prognosis, diagnosis, TMB, RNA modification, MSI, and immune infiltration. A reduction in CCDC51 expression was an independent prognostic factor in the TCGA-LIHC cohort and was directly related to improved OS, DSS, and PFS. CCDC51 can therefore be used to assess the prognosis of cancer patients and as a therapeutic target for LIHC diagnosis.

Data Availability

The data used to support the findings of this study are included within the article. The data and materials in the current study are available from the corresponding author on reasonable request.

Conflicts of Interest

The authors declare no conflicts of interest.

Author contributions

All authors participated in the interpretation of the studies, analysis of the data, and review of the manuscript. XYH, and SYW contributed to the study conception and design. SYW, PC, ZC contributed to the data analysis. The first draft of the manuscript was written by SYW, and XYH. All authors read and approved the final manuscript.

Funding

No funding was used in this study.

Acknowledgments

We thank the support of Home for Research (www.home-for-researchers.com).

References

[1] Siegel RL, Miller KD, Fuchs HE, Jemal A. Cancer Statistics, 2021. CA Cancer J Clin. 2021;71(1):7-33.

[2] Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, et al. Cancer statistics in China, 2015. CA Cancer J Clin. 2016;66(2):115-32.

[3] Zhang X, Li C, Wen T, Yan L, Li B, Yang J, et al. Appropriate treatment strategies for intrahepatic recurrence after curative resection of hepatocellular carcinoma initially within the Milan criteria: according to the recurrence pattern. Eur J Gastroenterol Hepatol. 2015;27(8):933-40.

[4] Anwanwan D, Singh SK, Singh S, Saikam V, Singh R. Challenges in liver cancer and possible treatment approaches. Biochim Biophys Acta Rev Cancer. 2020;1873(1):188314.

[5] Zeitz C, Méjécase C, Michiels C, Condroyer C, Wohlschlegel J, Foussard M, et al. Mutated CCDC51 Coding for a Mitochondrial Protein, MITOK Is a Candidate Gene Defect for Autosomal Recessive Rod-Cone Dystrophy. Int J Mol Sci. 2021;22(15).

[6] Li T, Fan J, Wang B, et al. TIMER: A Web Server for Comprehensive Analysis of Tumor-Infiltrating Immune Cells. Cancer Res. 2017. 77(21): e108-e110.

[7] Chandrashekar DS, Bashel B, Balasubramanya S, et al. UALCAN: A Portal for Facilitating Tumor Subgroup Gene Expression and Survival Analyses. Neoplasia. 2017. 19(8): 649-658.

[8] Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017. 45(W1): W98-W102.

[9] Mizuno H, Kitada K, Nakai K, Sarai A. PrognoScan: a new database for meta-analysis of the prognostic value of genes. BMC Med Genomics. 2009. 2: 18.

[10] Cerami E, Gao J, Dogrusoz U, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. Cancer Discov. 2012. 2(5): 401-4.

[11] Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, et al. Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. Cell. 2018;173(2):338-54.e15.

[12] Yang W, Soares J, Greninger P, et al. Genomics of Drug Sensitivity in Cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells. Nucleic Acids Res. 2013. 41(Database issue): D955-61.

[13] Basu A, Bodycombe NE, Cheah JH, et al. An interactive resource to identify cancer genetic and lineage dependencies targeted by small molecules. Cell. 2013. 154(5): 1151-1161.

[14] Addeo A, Friedlaender A, Banna GL, Weiss GJ. TMB or not TMB as a biomarker: That is the question. Crit Rev Oncol Hematol. 2021;163:103374.

[15] Luchini C, Bibeau F, Ligtenberg M, Singh N, Nottegar A, Bosse T, et al. ESMO recommendations on microsatellite instability testing for immunotherapy in cancer, and its relationship with PD-1/PD-L1 expression and tumour mutational burden: a systematic review-based approach. Ann Oncol. 2019;30(8):1232-43.

[16] Schrock AB, Ouyang C, Sandhu J, Sokol E, Jin D, Ross JS, et al. Tumor mutational burden

is predictive of response to immune checkpoint inhibitors in MSI-high metastatic colorectal cancer. Ann Oncol. 2019;30(7):1096-103.

[17] Boland CR, Goel A. Microsatellite instability in colorectal cancer. Gastroenterology. 2010;138(6):2073-87.e3.

[18] Gajewski TF, Schreiber H, Fu YX. Innate and adaptive immune cells in the tumor microenvironment. Nat Immunol. 2013;14(10):1014-22.

[19] Barbieri I, Kouzarides T. Role of RNA modifications in cancer. Nat Rev Cancer. 2020;20(6):303-22.

[20] Dai F, Wu Y, Lu Y, An C, Zheng X, Dai L, et al. Crosstalk between RNA m(6)A Modification and Non-coding RNA Contributes to Cancer Growth and Progression. Mol Ther Nucleic Acids. 2020;22:62-71.

[21] Dohlman AB, Arguijo Mendoza D, Ding S, Gao M, Dressman H, Iliev ID, et al. The cancer microbiome atlas: a pan-cancer comparative analysis to distinguish tissue-resident microbiota from contaminants. Cell Host Microbe. 2021;29(2):281-98.e5.

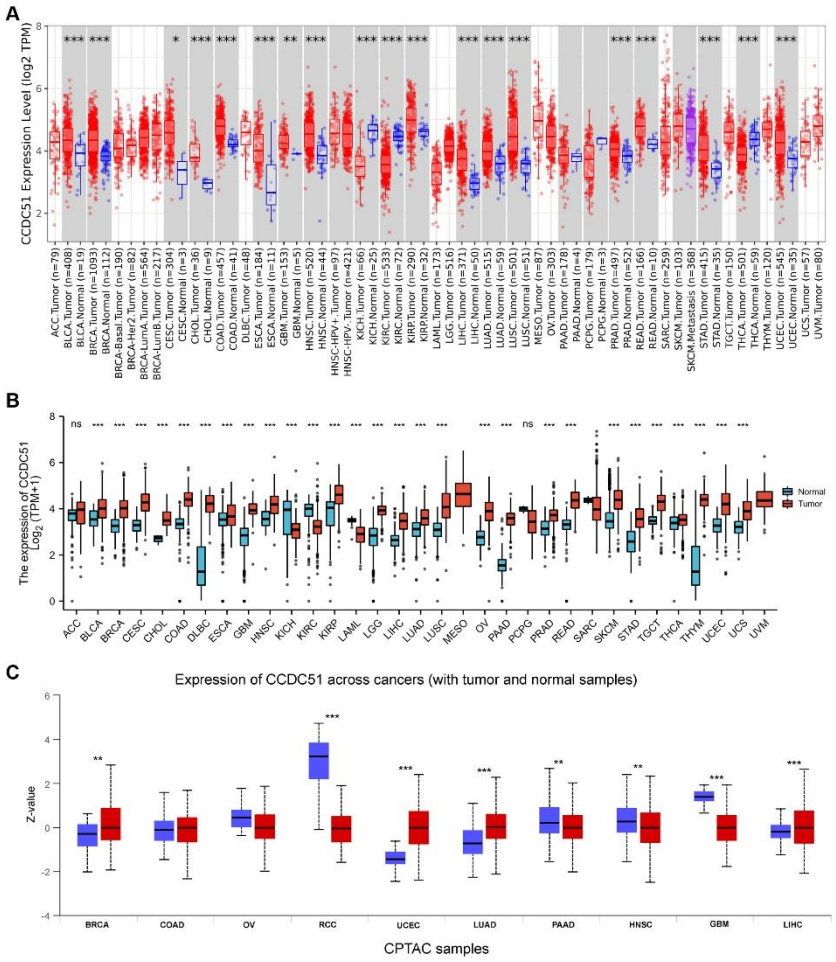


FIGURE 1. Tumors and normal tissues express CCDC51 differently.

(A) TIMER database analysis of CCDC51 expression in pan-cancer. (B) The pan-cancer expression of CCDC51 from the TCGA/GTEx database. (C) A pan-cancer analysis of CCDC51 by UALCAN.

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. ns, $p > 0.05$, *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

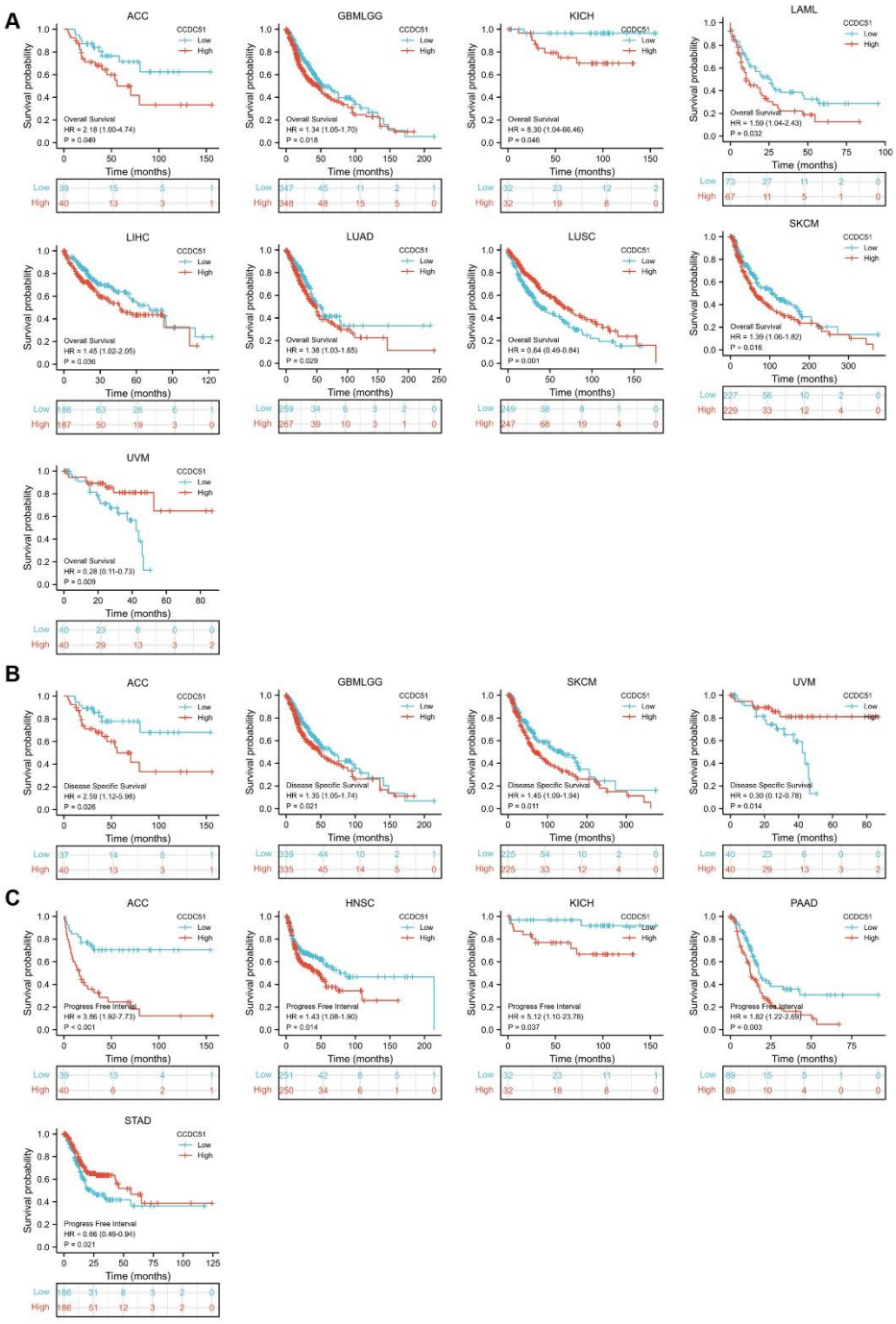


FIGURE 2. Prognosis values for pan-cancer were correlated with CCDC51 expression.

(A) Survival for CCDC51 in ACC, GBMLGG, KICH, LAML, LIHC, LUAD, LUSC, SKCM, and UVM. (B) CCDC51 disease-specific survival in ACC, GBMLGG, SKCM, and UVM. (C) The progress free interval for CCDC51 in ACC, HNSC, KICH, and PAAD.

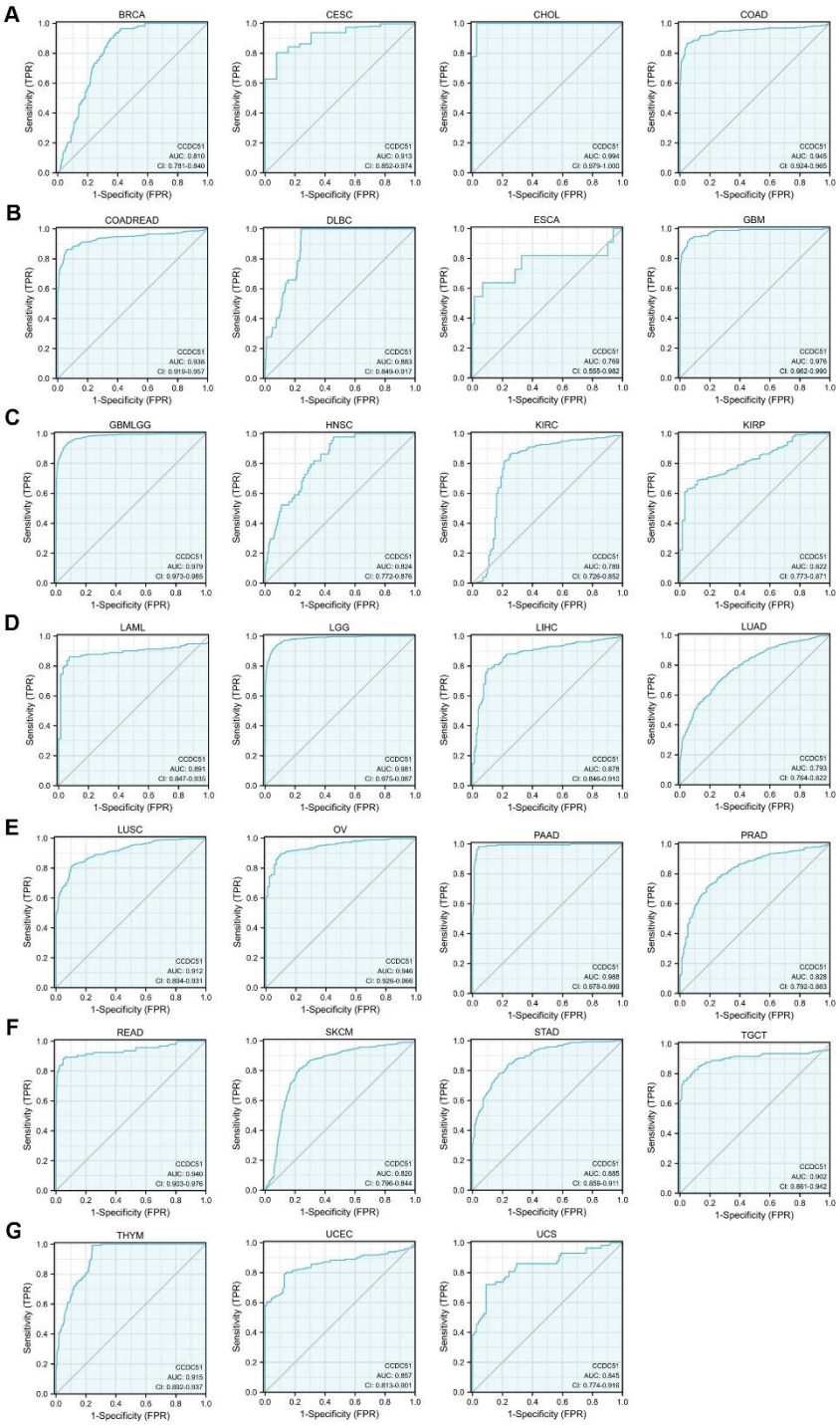


FIGURE 3. A potential biomarker for human cancer may be CCDC51.

A ROC curve analysis of CCDC51 expression predicts prognosis in BRCA, CESC, CHOL, and COAD (A); COADREAD, DLBC, ESCA, and GBM (B); GBMLGG, HNSC, KIRC, and KIRP (C); LAML, LGG, LIHC, and LUAD (D); LUSC, OV, PAAD, and PRAD (E); READ, SKCM, STAD, and TGCT (F); THYM, UCEC, and UCS (G).

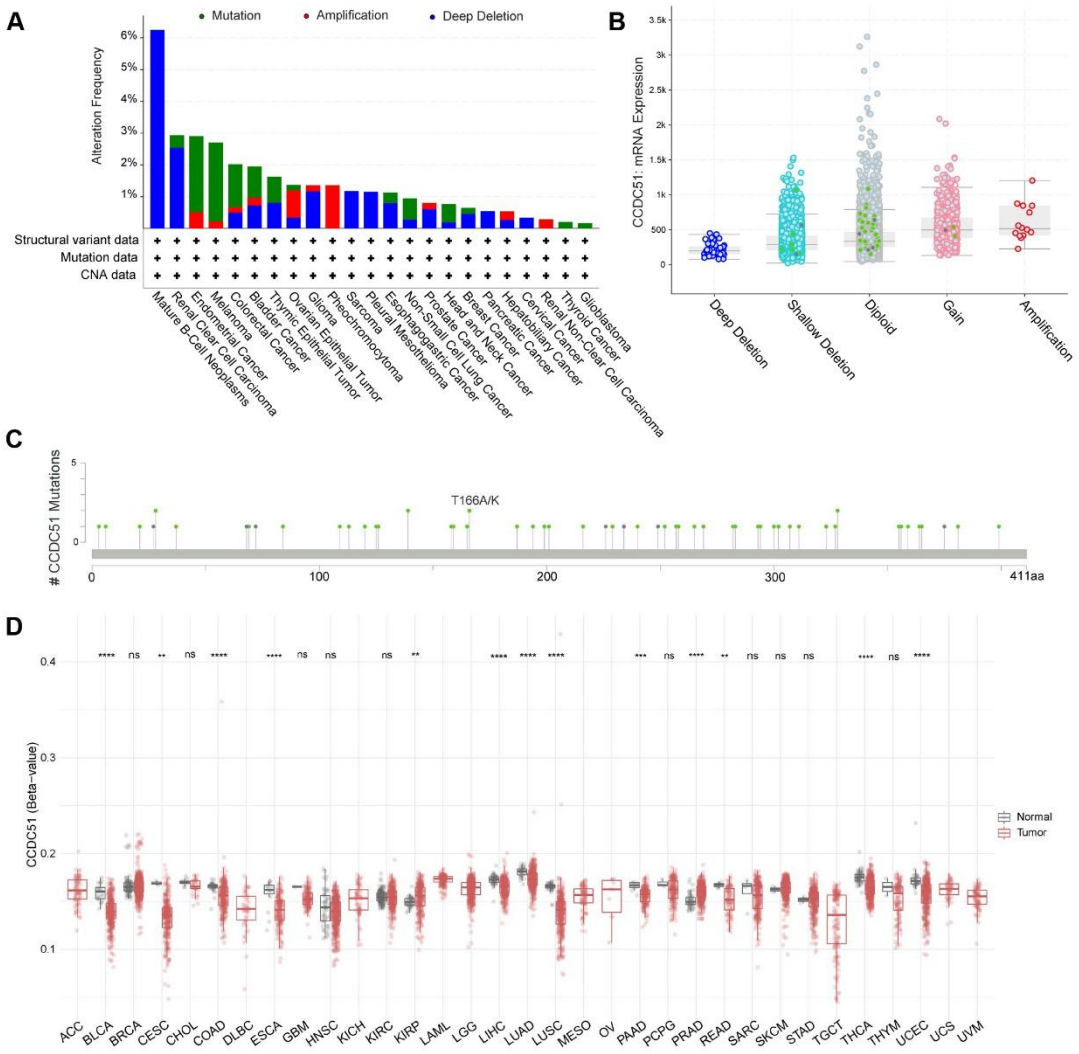


FIGURE 4. Mutational and DNA methylation analysis of CCDC51.

(A) Summary of mutation types of CCDC51 and the distribution among different cancers. (B) The cBioPortal database examined the frequency of mutation in CCDC51 in pan-cancer. (C) Hot spots of mutation of CCDC51. (D) Correlation between CCDC51 expression and DNA methylation.

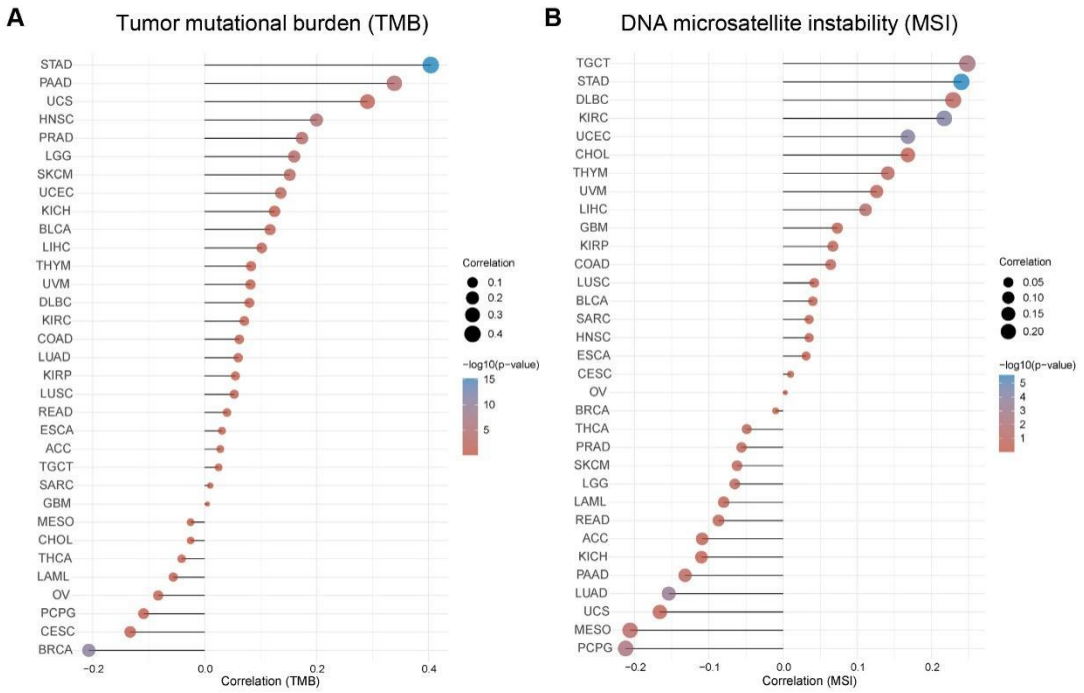


FIGURE 5. The relationship between CCDC51 expression and TMB and MSI. (A) Correlation between CCDC51 and TMS in pan-cancer. (B) Correlation between CCDC51 and MSI in pan-cancer.

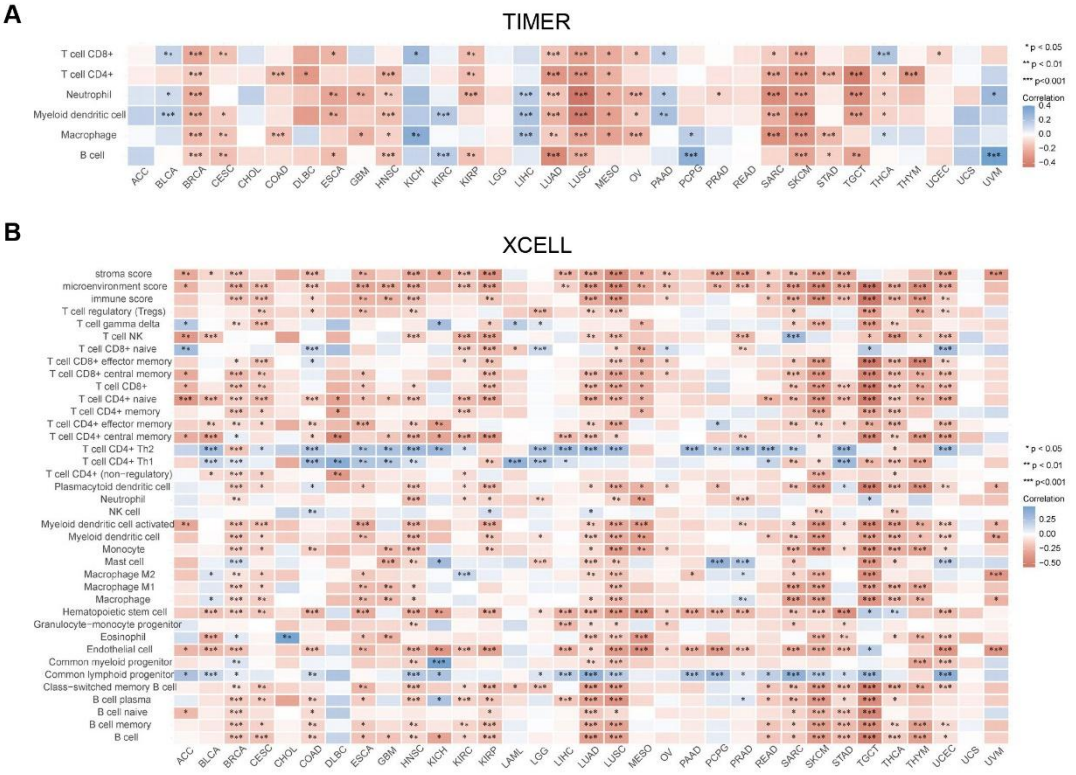


FIGURE 6. Relationship between CCDC51 expression and immune infiltration. (A) The correlations between CCDC51 expression and immunoinvasive level in 33 human tumors using TIMER. (B) Correlation between CCDC51 expression and immune invasion levels in 33 human tumors using xCell. *p <0.05, **p < 0.01, ***p < 0.001.

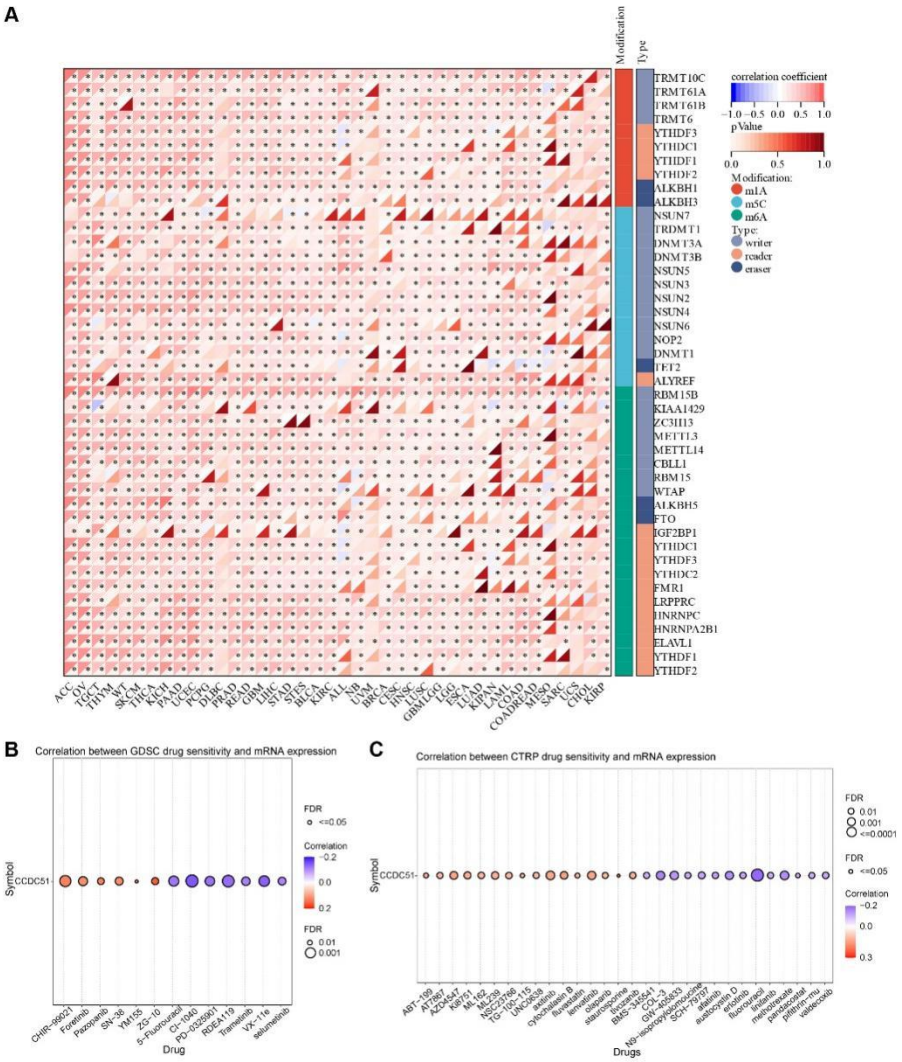
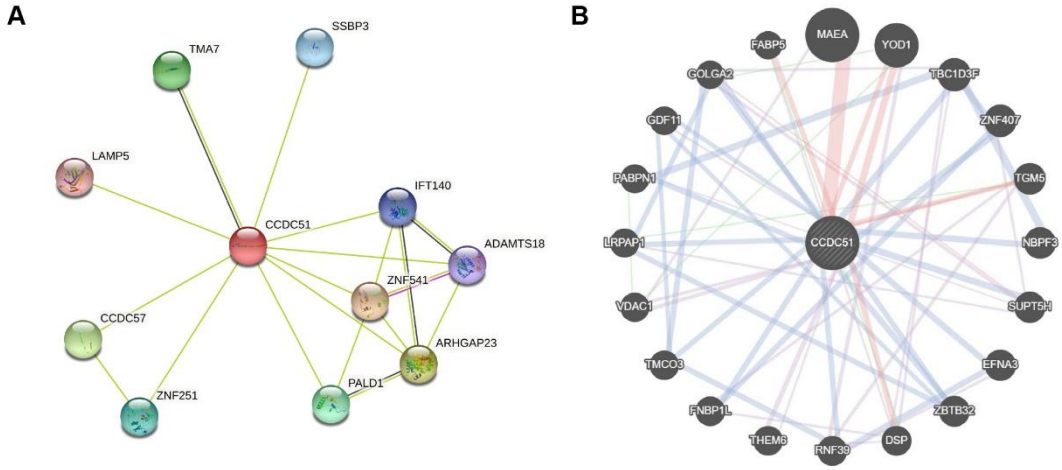


FIGURE 7. An analysis of the correlation between CCDC51 expression and RNA modification, as well as drug sensitivity in pan-cancer.

(A) Correlations between CCDC51 expression and and RNA modification. (B) Analyzing drug sensitivity in diverse human cancers using the GDSC database. (C) CTRP database analysis of CCDC51 expression and drug sensitivity in different human cancers. *$p < 0.05$.



410

411

412 **FIGURE 8. Interaction network of CCDC51 at the gene and protein levels**.

413 (A) Protein–protein interaction network of individual CCDC51 (STRING database).

414 (B) Gene–gene interaction network of individual CCDC51 (GeneMANIA database).

415

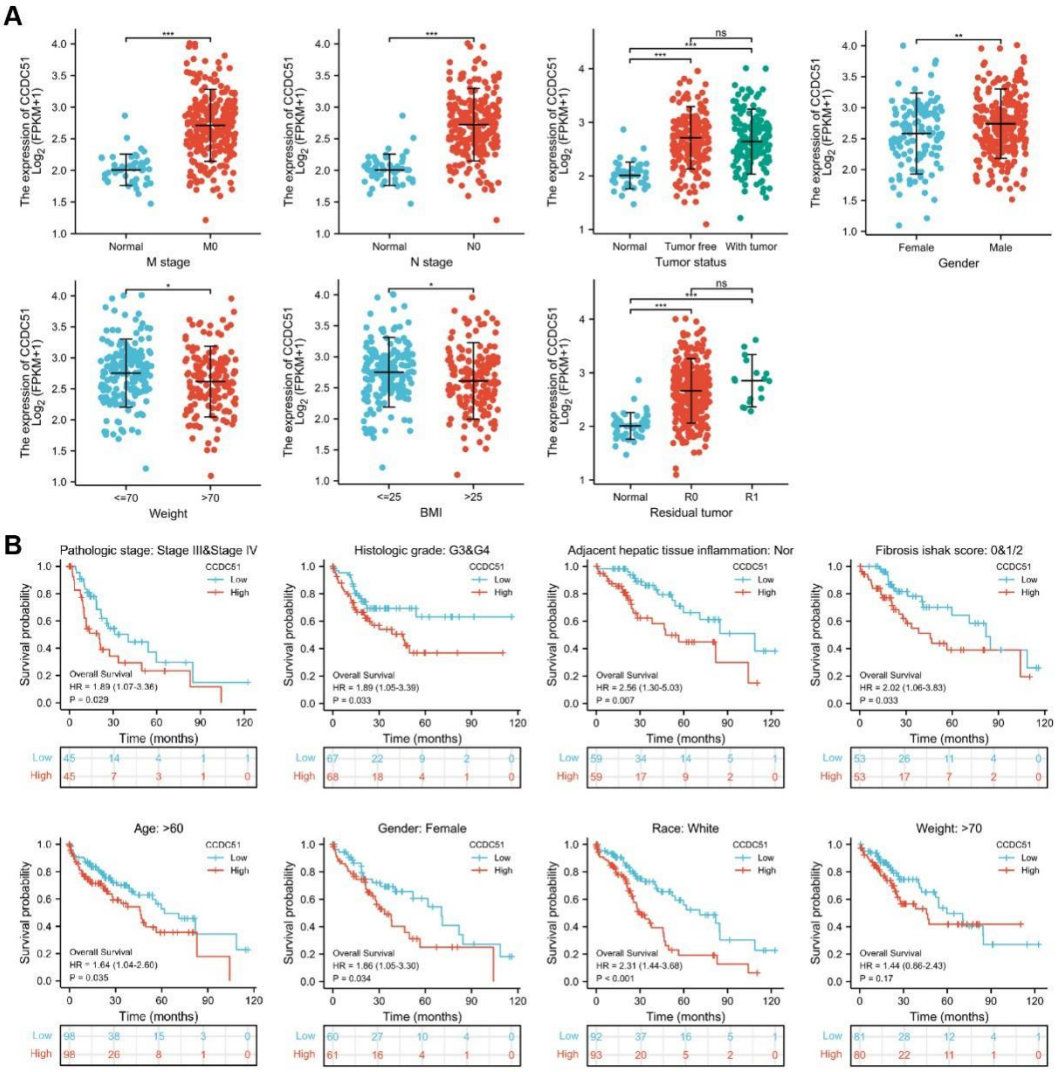


FIGURE 9. Prognostic analysis of CCDC51 in LIH.

(A) The correlation between CCDC51 expression and Clinical Characteristics in LIHC. (B) The correlation between CCDC51 and OS in different clinical subgroups of LIHC.

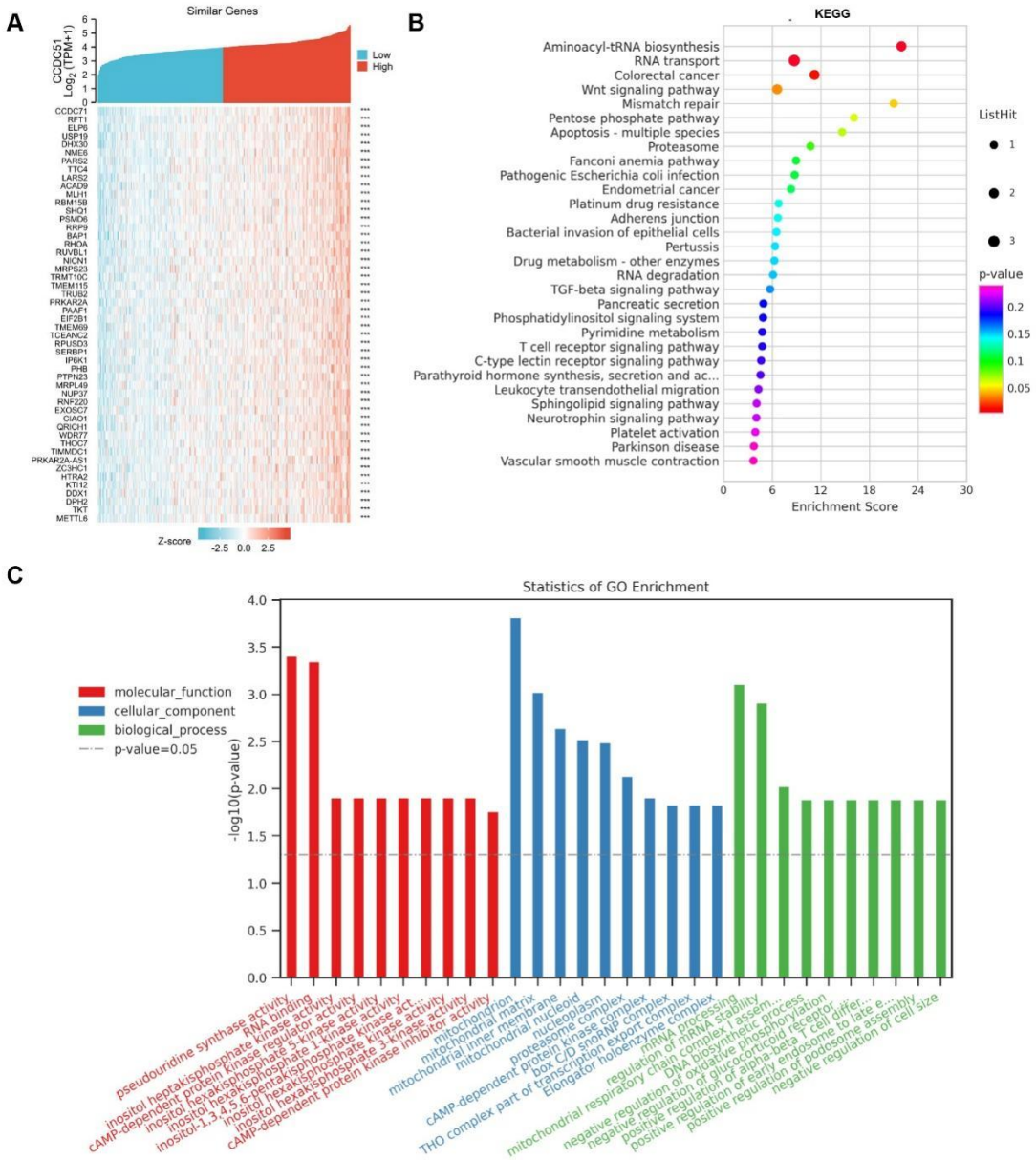


FIGURE 10. CCDC51 functional enrichment analysis in LIHC.

(A) Hot map of the top 50 similar genes with CCDC51 in LIHC. (B) KEGG of CCDC51 analysis by using the first 50 similar genes. (C) The GO term of CCDC51 analysis by using the first 50 similar genes.

Table 1. Univariate and multivariate Cox regression analyses of different parameters on OS in LIHC.

Characteristics	Total(N)	Univariate analysis		Multivariate analysis	
		Hazard ratio (95% CI)	P value	Hazard ratio (95% CI)	P value
T stage	370				
T1&T2	277	Reference			
T3&T4	93	2.598 (1.826-3.697)	**<0.001**	1.906 (0.255-14.267)	0.530
N stage	258				
N0	254	Reference			
N1	4	2.029 (0.497-8.281)	0.324		
M stage	272				
M0	268	Reference			
M1	4	4.077 (1.281-12.973)	**0.017**	1.500 (0.354-6.349)	0.582
Pathologic stage	349				
Stage I&Stage II	259	Reference			
Stage III&Stage IV	90	2.504 (1.727-3.631)	**<0.001**	1.270 (0.170-9.483)	0.816
Tumor status	354				
Tumor free	202	Reference			
With tumor	152	2.317 (1.590-3.376)	**<0.001**	1.959 (1.228-3.124)	**0.005**
CCDC51	373	1.826 (1.265-2.636)	**0.001**	1.858 (1.148-3.006)	**0.012**

Table 2. Univariate and multivariate Cox regression analyses of different parameters on DSS in LIHC.

Characteristics	Total(N)	Univariate analysis		Multivariate analysis	
		Hazard ratio (95% CI)	P value	Hazard ratio (95% CI)	P value
T stage	362				
T1&T2	272	Reference			
T3&T4	90	3.639 (2.328-5.688)	**<0.001**	14.664 (0.831-258.901)	0.067
N stage	253				
N0	249	Reference			
N1	4	3.612 (0.870-14.991)	0.077	9.936 (1.242-79.495)	**0.030**
M stage	268				
M0	265	Reference			
M1	3	5.166 (1.246-21.430)	**0.024**	2.319 (0.533-10.081)	0.262
Pathologic stage	341				
Stage I&Stage II	254	Reference			
Stage III&Stage IV	87	3.803 (2.342-6.176)	**<0.001**	0.279 (0.015-5.202)	0.392
Tumor status	354				
Tumor free	202	Reference			
With tumor	152	775790759.389 (0.000-Inf)	0.994		
CCDC51	365	1.768 (1.101-2.840)	**0.018**	1.842 (0.893-3.797)	0.098

Table 3. Univariate and multivariate Cox regression analysis of the influence of different parameters on PFI in LIHC patients.

Characteristics	Total(N)	Univariate analysis		Multivariate analysis	
		Hazard ratio (95% CI)	P value	Hazard ratio (95% CI)	P value
T stage	370				
T1&T2	277	Reference			
T3&T4	93	2.177 (1.590-2.980)	**<0.001**	0.864 (0.207-3.605)	0.841
N stage	258				
N0	254	Reference			
N1	4	1.370 (0.338-5.552)	0.659		
M stage	272				
M0	268	Reference			
M1	4	3.476 (1.091-11.076)	**0.035**	1.471 (0.451-4.804)	0.522
Pathologic stage	349				
Stage I&Stage II	259	Reference			
Stage III&Stage IV	90	2.201 (1.591-3.046)	**<0.001**	2.044 (0.489-8.548)	0.327
Tumor status	354				
Tumor free	202	Reference			
With tumor	152	11.342 (7.567-17.000)	**<0.001**	15.329 (9.216-25.497)	**<0.001**
CCDC51	373	1.234 (0.904-1.686)	0.186		