CCNA2 Expression and Its Prognostic Significance in cholangiocarcinoma

Jie Zhang
Clinical Medical College and The First Affiliated Hospital of Chengdu Medical College
Jingjun Zhang
The People's Hospital of JianYang City
Hairong Liu
Clinical Medical College and The First Affiliated Hospital of Chengdu Medical College
Qinyang Chen
Chengdu Medical College
Pengcheng Li
Southwest Medical University School of Clinical Medical Sciences
Taiyu Chen
Chengdu Medical College
Lei Xu
Clinical Medical College and The First Affiliated Hospital of Chengdu Medical College
Tao Ren (rentao509@cmc.edu.cn)
The First Affiliated Hospital of Chengdu Medical College

Research article

Keywords: cholangiocarcinoma, CCNA2, immune infiltration, prognosis

Posted Date: November 20th, 2020

DOI: https://doi.org/10.21203/rs.3.rs-111136/v1

Abstract

Background: Cholangiocarcinoma（CCA）is a rare malignancy and it has become a significant health burden worldwide. An increasing number of studies have demonstrated the crucial correlation of immunophenotypic characteristics modifications in resected specimen of CCA. However, the accurate prognostic markers is still lacking in the prognosis of CCA.

Methods: Gene expression profiles and clinical data of CCA were downloaded from the Gene Expression Omnibus（GEO）database. GO and KEGG analysis were applied for differentially expressed genes in cholangiocarcinoma, and PPI network was constructed in Cytoscape software. The expression difference of Cyclin A2 (CCNA2) in CCA tissues and adjacent noncancerous tissues was analyzed by R software and verified by comprehensive analysis. The relationship between CCNA2 expression and immune infiltration was assessed in the the Cancer Genome Atlas (TCGA) database. Kaplan–Meier survival analysis were chosen to assess the effect of CCNA2 expression on survival. Gene set enrichment analysis (GSEA) was used to screen the signaling pathways involved in CCA between the low and the high CCNA2 expression group.

Results: The expression of CCNA2 in CCA was significantly higher than that in adjacent cancerous tissues ($P < 0.001$) from the GEO database. The top 10 hub genes were mined by the Degree, MCC, and Closeness method based on the PPI network. CCNA2 was screened as the candidates to be further analyzed and validated. The Kaplan–Meier curves suggested that patients with high CCNA2 expression had a poor prognosis. Multivariate analysis showed that a high expression of CCNA2 was an important independent predictor of poor overall survival ($P = 0.033$). Cibersort analysis showed that the fraction of T cells CD4 (naive, $P < 0.0001$), T cells CD4 (memory, $P = 0.0055$), T cells (follicular, $P = 0.0063$), NK cells ($P = 0.0149$), dendritic cells ($P = 0.029$), neutrophils ($P = 0.0113$) is significantly correlated with the expression of CCNA2, which highlighted the CCNA2 expression in immune infiltrates. GSEA indicated that 12 signaling pathways were evidently enriched in samples with the high-CCNA2 phenotype.

Conclusions: CCNA2 might act as an oncogene in the progression of CCA and could be regarded as a potential prognostic indicator and therapeutic target for CCA.

1. Introduction

Cholangiocarcinoma (CCA) is the most common malignant tumor in the intrahepatic biliary epithelium. Its incidence accounts for about 3% of all gastrointestinal tumours and is the second most common primary liver tumour after hepatocellular carcinoma [1]. The onset of CCA is concealed, with no obvious symptoms or signs in the early stages [2]. Delayed clinical diagnosis limits the benefit of surgical treatment and curative management options, contributing to the poor outcome of CCA patients [3]. At present, surgical resection or liver transplantation is the main treatment method for CCA, but the prognosis of patients with surgical resections is poor, and the availability of livers for liver transplantation

is rare. The long-term survival rate is still low, which causes great pain and economic burden to patients and their families [4]. Therefore, the early detection of CCA is essential for providing patients a curative therapeutic approach and to maximize clinical outcomes, prognostic biomarkers might help guide treatment decisions with respect to patient life expectancy and develop personalized treatments for individual CCA patients.

High-throughput genomics and epigenomics have greatly increased our understanding of CCA underlying biology, however its pathogenesis remains largely unknown. CCA is characterized by a highly desmoplastic microenvironment containing stromal cells, mainly cancer-associated fibroblasts, infiltrating tumor epithelium [5]. Tumor microenvironment in CCA is a highly dynamic environment that, besides stromal and endothelial cells, encompass also an abundance of immune cells, of both the innate and adaptive immune system (including tumor-associated macrophages, neutrophils, and T and B lymphocytes) and abundant proliferative factors[6, 7]. It is orchestrated by multiple soluble factors and signals, that eventually define a tumor growth-permissive microenvironment [8, 9]. Through complicate interactions with CCA cells, tumor microenvironment profoundly affects the proliferative and invasive abilities of epithelial cancer cells and plays an important role in accelerating neovascularization and preventing apoptosis of neoplastic cells[10].

The present study combined bioinformatics, transcriptomics and epidemiology in order to solve this issue and explore the molecular mechanism of CCA. Based on The Cancer Genome Atlas (TCGA) dataset, the Gene Expression Omnibus (GEO) dataset, we try to create a comprehensive “atlas” of CCA, and discover major cancer-causing genomic alterations [11]. In addition to their expression, their associated regulatory network was investigated. The data can elucidate the underlying etiology of CCA and provide reliable molecular targets for drug therapy. Further we discuss innovative treatment approaches, including immunotherapy, and how identification of CCA secreted factors by immune cell subsets are leading towards a precision medicine in CCA.

2. Results

2.1. Identification of DEGs in cholangiocarcinoma (CCA)

We collected two datasets from GEO, including *GSE26566* and *GSE32225* and used SVA package to remove batch effects and other unwanted variation in experiments and obtained an integrated GEO datasets. Then, limma packages of R were used to screen for DEGs in the integrated GEO datasets. Finally, a total of 588 DEGs were identified, comprising 216 upregulated genes and 372 downregulated genes (Fig. 1A) at a cutoff |log2 fold change| >1 and padj < 0.05 (p-value adjusted for multiple testing using Benjamini-Hochberg method). The volcano plot of the DEGs is shown in Fig. 1B.

Figure 1. Identification of differentially expressed genes in cholangiocarcinoma (CCA) versus adjacent noncancerous tissue. (A) Heat map showing the expression profiles of mRNAs. X-axis represents

samples, Y-axis represents genes, red stands for upregulation, and blue stands for downregulation (B) Significant DEGs in CCA versus adjacent noncancerous tissue.

3.2 *Functional enrichment analysis*

GO and KEGG analyses were performed to examine the biological functions of DEGs in detail [12]. The results showed that complement and coagulation cascades, chemical carcinogenesis, and teroid hormone biosynthesis were significantly enriched (Fig. 2A.B). GO analysis also described the results from 3 categories: monocarboxylic acid metabolic process,organic hydroxy compound metabolic process□gland development. (Figure 2A). The bubble plot offered a visual representation of the aforementioned pathways.

3.3. *PPI network construction and identification of hub genes*

The PPI network analysis was further used to study the molecular mechanism of diseases and identify novel drug targets from a systematic perspective. The PPI network was conducted by STRING to explore the interactions between proteins encoded by these DEGs. We constructed a network of protein-protein interaction (PPI) by Cytoscape consisted of 447 nodes and 1301 edges according to the STRING database (Fig. 3A). Moreover, CytoHubba provides a user-friendly interface to examine the interactions among hub nodes in the biological network by topological analysis, which can aid the identification of the essential networks involved. The top 10 hub genes were mined by three network topology parameters, including Degree, MCC, and Closeness (Fig. 3B-D). Finally, CCNA2 was screened as the candidates to be further analyzed and validated (Fig. 3E).

Table 1
Top 10 hub genes according to three network topology parameters, including Degree, MCC, and Closeness.

Topology parameters	Hub genes
Degree	TOP2A,CCNA2,MAD2L1,G6PD,RRM2,CDKN2A,CYCS,EZH2,COL1A1,CXCL8
MCC	TOP2A,CCNA2,MAD2L1,UHRF1,CENPF,RRM2,KIAA0101,NCAPG,NUSAP1,CDCA5
Closeness	CCNA2,SMAD3,G6PD,CDKN2A,PKM,CYCS,SPP1,EZH2,CXCL8,NOS3

3.4 The Difference in Cyclin A2 (CCNA2) expression in CCA

The Cyclin A2 (CCNA2) expression data at the mRNA level were obtained from the GEO database (including 253 CCA tissues and 12 adjacent non-tumor tissues) and the TCGA database (including 36 CCA tissues and 9 adjacent non-tumor tissues). We found that the expression of CCNA2 upregulated in CCA tissues compared with that in adjacent noncancerous tissues, but especially in TCGA database (P< 0.0001). The scatter plot shows the mRNA expression profiles of CCNA2 in CCA tissues and adjacent noncancerous tissues.

3.5 High Expression of CCNA2 in CCA Is Related to Poor Overall Survival

We evaluated the prognosis of high-CCNA2 expression in CCA patients from the TCGA by Kaplan–Meier risk estimates. The results revealed that, compared with the low CCNA2 expression, the high CCNA2 expression was more significantly associated with a poor overall survival ($P = 0.033$, Fig. 5). The median OS of the high-CCNA2 expression group was 21 months while the median OS of the low-CCNA2 expression group was 31 months. As depicted in Fig. 5, the overexpression of CCNA2 genes is significantly correlated with poor prognosis in CCA patients ($p < .05$).

3.5 *The Co-expression Analysis*

According to CIBERSORT algorithm, the proportion of 22 types of tumor-infiltrating immune cells in each sample of integrated GEO datasets were calculated, and 187 samples with a CIBERSORT P- value of < 0.05 were included in the next analysis. Then, co-expression analysis was did between proportions of immune cells and the expression of CCNA2. The results showed that the fraction of T cells CD4 (naive, $P < 0.0001$), T cells CD4 (memory, $P = 0.0055$), T cells (follicular, $P = 0.0063$), NK cells ($P = 0.0149$), dendritic cells ($P = 0.029$), neutrophils ($P = 0.0113$) is significantly associated with the expression of CCNA2, respectively (Figs. 6A–F), which showed CCNA2 plays an important role in tumor immune microenvironment.

3.6 *Identification of CCNA2-Related Signaling Pathways by GSEA*

On the basis of the KEGG and GO data, we explored the function of CCNA2 and its related signal transduction pathway through GSEA. In view of NES, FDR *q*-value, and nominal *p*-value, significantly enriched signaling pathways were selected. In this study, 12 signaling pathways involved in cytokine-cytokine receptor interaction,complement and coagulation cascades, natural killer cell mediated cytotoxicity, cell cycle, ribosome,and calcium signaling pathway were differentially enriched in the highly expressed phenotypes of CCNA2 (Table 2–3, Fig. 7).

Table 2
Gene sets enriched in the high CCNA2 expression (based on KEGG data).

ID	Description	Enrichment Score	NES	P value
hsa04610	Complement and coagulation cascades	-0.653	-2.394	0.002
hsa05204	Chemical carcinogenesis	-0.543	-1.934	0.002
hsa00140	Steroid hormone biosynthesis	-0.548	-1.859	0.002
hsa00982	Drug metabolism - cytochrome P450	-0.580	-2.010	0.002
hsa00830	Retinol metabolism	-0.594	-2.032	0.002
hsa00980	Metabolism of xenobiotics by cytochrome P450	-0.531	-1.844	0.002

Table 3
Gene sets enriched in the high CCNA2 expression (based on GO data).

ID	Description	Enrichment Score	NES	P value
GO:0032787	monocarboxylic acid metabolic process	-0.377	-1.702	0.001
GO:1901615	organic hydroxy compound metabolic process	-0.342	-1.528	0.001
GO:0048732	gland development	-0.341	-1.513	0.001
GO:0043588	skin development	-0.365	-1.601	0.001
GO:0008544	epidermis development	-0.365	-1.607	0.001
GO:0044282	small molecule catabolic process	-0.389	-1.707	0.001

3. Discussion

Cell cycle progression is regulated by the interaction of cyclin, cyclin dependent kinases [13]. Therefore, the detection of their expression are important in progression and control of the cell cycle in many tumours [14]. Several studies demonstrated high cyclin B1 expression in different carcinomas, which was associated with aggressiveness of tumour, and might serve as a prognostic marker [15]. It can be present in the cytoplasm and in the nucleus, the latter associated with poorer prognosis (as demonstrated in breast and oesophageal carcinomas) [16, 17]. Furthermore, the cyclin A2 (CCNA2)-CDK2 complex is essential for the progression of G2 phase into mitosis [18]. CCNA2 is localised in the nucleus during S phase, controlling DNA synthesis [19].

In this study, we sought to determine the role of CCNA2 expression in CCA progression, especially as a prognostic factor for CCA. In addition, we also tried to screen signaling pathways related to CCNA2 in

CCA to understand the underlying mechanism involved in the regulation of CCA development by CCNA2. First, we analyzed the data in the GEO database and compared the differentially expressed genes in CCA and adjacent noncancerous tissues. Then, the PPI network analysis was used to further screen real hub genes with a significant *p* value. It's worth noting that CCNA2 were especially outstanding. It was not only highly correlated with CCA grade, but also may be potential biomarkers for prognosis. It is noteworthy that we established Kaplan–Meier risk estimates to predict survival of CCA patients based on tumor-infiltrating immune cells, the survival rate of patients in the low-CCNA2 expression group was higher than that of patients in the high-CCNA2 expression group.

The tumor microenvironment often affects the invasive processes. The extracellular matrix molecules and secreted growth factors are involved in the transition of tumor cells into an invasive phenotype. The invasion and metastasis of tumor cells may have nothing to do with the proliferation of tumor cells but have occurred already at the early developmental stage of the tumor [20]. Previous studies about tumors and immune infiltration mostly focused on immune cell types. For example, Li B *et al.* infers the abundance of the six immune cell types (B cells, CD4 T cells, CD8 T cells, neutrophil, macrophage, and dendritic cells) using approach of constrained least squares fitting and found many significant associations between immune cell abundance and outcome of 23 cancer types patients [21]. For instance, except for association with prolonged survival of patients, CD8 T cells may also play an important role in preventing tumor recurrence (in melanoma and colorectal cancer and cervical cancer. Notably, the focus of our research is to estimate the degree of 6 immune cells infiltration, and finally determined their most promising co-expression patterns associated with CCNA2.These differences may illustrate that CCNA2 may affect the immune microenvironment of CCA to a certain extent.

Based on that, in this study, GSEA was implemented to calculate the immune cell infiltration levels for each sample. Since previous studies had shown immune-infiltration to have better prognosis in different carcinomas, we analyzed the CCNA2 expression in differential groups. CCA-related tumor samples from various databases and literatures were collected. Functional analysis showed these samples to be closely associated with the expression of CCNA2, such as via cytokine-cytokine receptor interaction,natural killer cell mediated cytotoxicity, cell cycle and calcium signaling pathway.

Our research also has some limitations. First, the clinical information is not perfect, and some important information, such as tumor size, was not provided. Second, there is a lack of specific details, such as surgical treatment and surgical details, which are crucial to the prognosis of patients. Finally, it is impossible to evaluate the protein level and direct mechanism of CCNA2 in CCA from GEO and TCGA database.

In conclusion, our study first analyzed the GEO and TCGA database and found that the expression of CCNA2 in CCA tissues is higher than that in adjacent noncancerous tissues. The upregulation of CCNA2 is closely correlated with some clinicopathological features of CCA, which are related to the occurrence and the development of CCA. Importantly, the fraction of 6 immune-related cells is significantly correlated with the expression of CCNA2, which highlighted the immunotherapy might actually change the

therapeutic landscape of CCA. In summary, we found that the expression level of CCNA2 may be a marker for the diagnosis and the prognosis of CCA. In future analyses, other clinical trials will be needed to verify the corresponding results to reveal the prognostic value of CCNA2 in CCA

4. Materials And Methods

4.1. *Data collection and processing*

We obtained the expressing profiling dataset GSE26566 and GSE32225 from the publicly available Gene Expression Omnibus database (GEO, https://www.ncbi.nlm.nih.gov/geo/*).* The expression dataset (GSE26566) contained 111 pairs of CCA tumors and matched non-tumor tissues, which was based on the GPL6104 platform (IlCCNA2ina HiSeq 2000). The expression dataset (GSE32225), which was generated on the GPL8432 platform (IlCCNA2ina HumanMethylation27 BeadChip), included a total of 155 primary CCA tissues and matched adjacent normal tissues (Table 4). GEOquery R package was used to download the expression profiles [22]. The SVA package was used to remove batch effects and other unwanted variation in experiments [23]. The differentially expressed genes (DEGs) of HNSCC were identified using “DESeq2” R package at a cutoff |log2 fold change| >1 and padj < 0.05 (p-value adjusted for multiple testing using Benjamini-Hochberg method).

Table 4
Summary of datasets and platforms

GSE	N tumor	N normal	Platform
GSE26566	104	6	GPL6104
GSE32225	149	6	GPL8432

4.2. *Functional enrichment analysis*

Gene ontology (GO) analysis provides a controlled vocabulary to describe gene and gene product attributes in any organism (http://www.geneontology.org*)* [24]. Pathway analysis is used to map genes to Kyoto Encyclopedia of Genes and Genomes (KEGG, https://www.kegg.jp/*).* In this paper, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enrichment analysis and Gene Ontology (GO) annotation were performed using the R package “clusterProfiler”. GO terms with P value < 0.05 and KEGG pathways with P value < 0.05 were considered to have significance.

4.3. *PPI network construction, module analysis, and identification of hub genes*

The PPI network of DEGs was constructed using the STRING online database with default parameters [25]. Cytoscape (http://www.cytoscape.org/*)* software was utilized to visualize the network [26]. Hub genes were identified using the CytoHubba according to three network topology parameters, including Degree, MCC, and Closeness [27].

4.4. *CCNA2 Expression Analysis and Survival Analysis*

Firstly, we extract the CCNA2 expression data from the integrated GEO datasets and TCGA dataset. Student's t-test was used to compare gene expression between tumor tissues and adjacent nontumorous tissues. Clinical information of patients in TCGA and CCNA2 expression data are combined into a matrix. According to the appropriate value of CCNA2 expression, the samples were divided into two groups (high-CCNA2 expression group and low-CCNA2 expression group). R software was applied to draw survival curves to visualize the impact of CCNA2 expression on patients' overall survival using survival package.

4.5 *CIBERSORT estimation*

We uploaded the gene expression data with standard annotation to the CIBERSORT web portal (http://cibersort.stanford.edu/*)*, and the algorithm was run using the LM22 signature and 1000 permutations [28]. Cases with a CIBERSORT output of $p < 0.05$, indicating that the inferred fractions of immune cell populations produced by CIBERSORT are accurate, were considered to be eligible for further analysis [29].

4.6 *Gene Set Enrichment Analysis*

GSEA was used to explore the signaling pathways related to CCNA2 in CCA. Gene expression enrichment analysis was carried out between datasets with low or high CCNA2 mRNA expression. The phenotype was determined by the expression level of CCNA2 based on the TCGA database. The annotated gene set was selected (c2.cp.kegg.v6.2.symbols.gmt) as the reference gene set. A total of 1,000 gene sets were arranged in each analysis to determine significantly different pathways. Gene set permutations were performed 1,000 times for each analysis to identify significantly different pathways. The normalized enrichment score (NES), nominal *p*-value, and false discovery rate (FDR) q-value indicated the importance of the association between gene sets and pathways.

4.7 *Statistical analysis*

The difference in CCNA2 expression between CCA tissues and adjacent noncancerous tissues was tested by Mann–Whitney U test. The Wilcoxon rank-sum test was a nonparametric statistical test mainly utilized for comparing two groups and the Kruskal–Wallis test was suitable when it comes to two or more groups. Kaplan–Meier analysis and log-rank test were used to compare the significant differences in survival rates between the high- and the low-CCNA2 expression groups. All statistical analyses were performed with IBM SPSS statistical software (version 23.0) and R software (version 2.15.3), and $P < 0.05$ was used to determine the significance level.

Declarations

We declare that we have no financial and personal relationships with other people or organizations that can inappropriately influence our work, there is no professional or other personal interest of any nature or kind in any product, service and/or company that could be construed as influencing the position

presented in, or the review of, the manuscript entitled “CCNA2 Expression and Its Prognostic Significance in cholangiocarcinoma”.

We confirm that we would like support with depositing and managing our data , and we can also submit datasets to Springer Nature as part of our publisher's Data Support Services.All data generated or analysed during this study are included in this published article.

Ethics approval and consent to participate There are no animal or human trials involved.

Consent for publication Not applicable.
Competing interests The authors declare that they have no competing interests.
Funding Not applicable
Authors' contributions All authors analyzed and interpreted the research data. Jie Zhang, Jingjun Zhang, Hairong Liu and Tao Ren was the major contributors in writing the manuscript.Meanwhile, all authors read and approved the final manuscript.

Acknowledgements We appreciated Clinical Medical College and The First Affiliated Hospital of Chengdu Medical College,The People's Hospital of JianYang City,Chengdu Medical College for supporting the study.

References

1. Rizvi S, Gores GJ. Pathogenesis, Diagnosis, and Management of Cholangiocarcinoma. Gastroenterology. 2013;145(6):1215–29.
2. Blechacz B, Komuta M, Roskams T, Gores GJ. Clinical diagnosis and staging of cholangiocarcinoma. Nature Reviews Gastroenterology Hepatology. 2011;8(9):512.
3. Nathan H, Pawlik TM, Wolfgang CL, Choti MA, Cameron JL, Schulick RD. Trends in Survival after Surgery for Cholangiocarcinoma: A 30-Year Population-Based SEER Database Analysis. J Gastrointest Surg. 2007;11(11):1488–97.
4. Khan SA, Taylor-Robinson SD, Toledano MB, Beck A, Elliott P, Thomas HC. Changing international trends in mortality rates for liver, biliary and pancreatic tumours. J Hepatol. 2002;37(6):806–13.
5. Nakamura H, Arai Y, Totoki Y, Shirota T, Elzawahry A, Kato M, Hama N, Hosoda F, Urushidate T, Ohashi S, Hiraoka N, Ojima H, Shimada K, Okusaka T, Kosuge T, Miyagawa S, Shibata T. Genomic spectra of biliary tract cancer. Nat Genet. 2015;47(9):1003–10.
6. Mao ZY, Zhu GQ, Xiong M, Ren L, Bai L. Prognostic value of neutrophil distribution in cholangiocarcinoma. *World journal of gastroenterology* 2015, 21, (16), 4961–8.
7. Morisaki T, Umebayashi M, Kiyota A, Koya N, Tanaka H, Onishi H, Katano M. Combining cetuximab with killer lymphocytes synergistically inhibits human cholangiocarcinoma cells in vitro. Anticancer research. 2012;32(6):2249–56.

8. Miura T, Yoshizawa T, Hirai H, Seino H, Morohashi S, Wu Y, Wakiya T, Kimura N, Kudo D, Ishido K, Toyoki Y, Kijima H, Hakamada K. Prognostic Impact of CD163 + Macrophages in Tumor Stroma and CD8 + T-Cells in Cancer Cell Nests in Invasive Extrahepatic Bile Duct Cancer. *Anticancer research* 2017, 37, (1), 183–190.
9. Oshikiri T, Miyamoto M, Shichinohe T, Suzuoki M, Hiraoka K, Nakakubo Y, Shinohara T, Itoh T, Kondo S, Katoh H. Prognostic value of intratumoral CD8 + T lymphocyte in extrahepatic bile duct carcinoma as essential immune response. Journal of surgical oncology. 2003;84(4):224–8.
10. Goeppert B, Frauenschuh L, Zucknick M, Stenzinger A, Andrulis M, Klauschen F, Joehrens K, Warth A, Renner M, Mehrabi A, Hafezi M, Thelen A, Schirmacher P, Weichert W. Prognostic impact of tumour-infiltrating immune cells on biliary tract cancer. British journal of cancer. 2013;109(10):2665–74.
11. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. *Contemporary oncology (Poznan, Poland)* 2015, 19, (1a), A68-77.
12. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology* 2012, 16, (5), 284–7.
13. Sherr CJ. The Pezcoller lecture: cancer cell cycles revisited. Cancer research. 2000;60(14):3689–95.
14. Musgrove EA, Caldon CE, Barraclough J, Stone A, Sutherland RL. Cyclin D as a therapeutic target in cancer. Nat Rev Cancer. 2011;11(8):558–72.
15. Koliadi A, Nilsson C, Holmqvist M, Holmberg L, de La Torre M, Wärnberg F, Fjällskog ML. Cyclin B is an immunohistochemical proliferation marker which can predict for breast cancer death in low-risk node negative breast cancer. Acta oncologica (Stockholm Sweden). 2010;49(6):816–20.
16. Nozoe T, Korenaga D, Kabashima A, Ohga T, Saeki H, Sugimachi K. Significance of cyclin B1 expression as an independent prognostic indicator of patients with squamous cell carcinoma of the esophagus. Clinical cancer research: an official journal of the American Association for Cancer Research. 2002;8(3):817–22.
17. Suzuki T, Urano T, Miki Y, Moriya T, Akahira J, Ishida T, Horie K, Inoue S, Sasano H. Nuclear cyclin B1 in human breast carcinoma as a potent prognostic factor. Cancer Sci. 2007;98(5):644–51.
18. De Boer L, Oakes V, Beamish H, Giles N, Stevens F, Somodevilla-Torres M, Desouza C, Gabrielli B. Cyclin A/cdk2 coordinates centrosomal and nuclear mitotic events. *Oncogene* **2008**, 27, (31), 4261–8.
19. Pagano M, Pepperkok R, Verde F, Ansorge W, Draetta G. Cyclin A is required at two points in the human cell cycle. EMBO J. 1992;11(3):961–71.
20. Hosseini H, Obradović MMS, Hoffmann M, Harper KL, Sosa MS, Werner-Klein M, Nanduri LK, Werno C, Ehrl C, Maneck M, Patwary N, Haunschild G, Gužvić M, Reimelt C, Grauvogl M, Eichner N, Weber F, Hartkopf AD, Taran FA, Brucker SY, Fehm T, Rack B, Buchholz S, Spang R, Meister G, Aguirre-Ghiso JA, Klein C. A., Early dissemination seeds metastasis in breast cancer. Nature. 2016;540(7634):552–8.
21. Li B, Severson E, Pignon JC, Zhao H, Li T, Novak J, Jiang P, Shen H, Aster JC, Rodig S, Signoretti S, Liu JS, Liu XS. Comprehensive analyses of tumor immunity: implications for cancer immunotherapy. Genome biology. 2016;17(1):174.

22. Davis S, Meltzer PS. GEOquery: a bridge between the Gene Expression Omnibus (GEO) and BioConductor. Bioinformatics. 2007;23(14):1846–7.
23. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics. 2012;28(6):882–3.
24. Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, Simonovic M, Doncheva NT, Morris JH, Bork P, Jensen LJ, Mering CV, STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic acids research* 2019, 47, (D1), D607-d613.
25. Szklarczyk D, Morris JH, Cook H, Kuhn M, Wyder S, Simonovic M, Santos A, Doncheva NT, Roth A, Bork P, Jensen LJ, von Mering C. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic acids research* **2017**, 45, (D1), D362-d368.
26. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome research. 2003;13(11):2498–504.
27. Chin CH, Chen SH, Wu HH, Ho CW, Ko MT, Lin CY. cytoHubba: identifying hub objects and sub-networks from complex interactome. BMC systems biology. 2014;8(Suppl 4):(Suppl 4), S11.
28. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. Nature methods. 2015;12(5):453–7.
29. Ali HR, Chlon L, Pharoah PD, Markowetz F, Caldas C. Patterns of Immune Infiltration in Breast Cancer and Their Clinical Implications: A Gene-Expression-Based Retrospective Study. PLoS Med. 2016;13(12):e1002194.

Figures

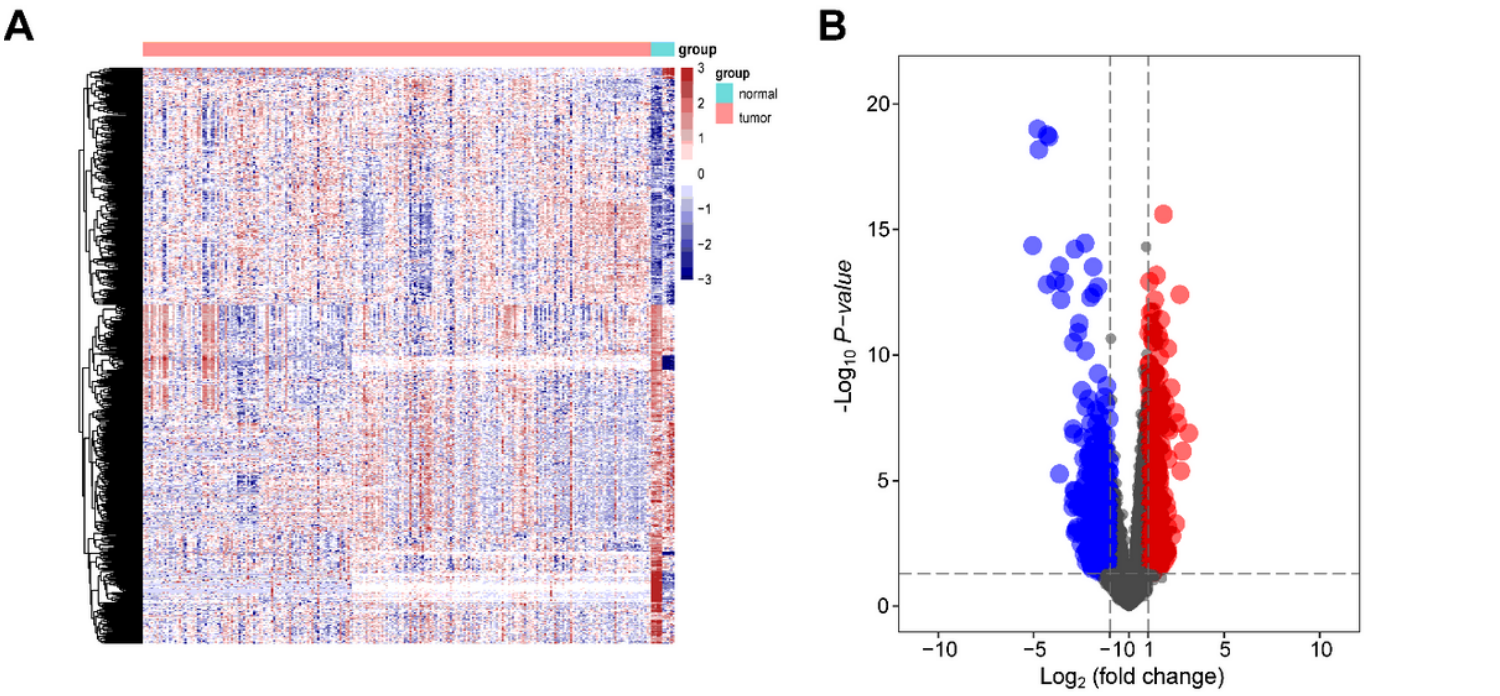

Figure 1

Identification of differentially expressed genes in cholangiocarcinoma (CCA) versus adjacent noncancerous tissue. (A) Heat map showing the expression profiles of mRNAs. X-axis represents samples, Y-axis represents genes, red stands for upregulation, and blue stands for downregulation (B) Significant DEGs in CCA versus adjacent noncancerous tissue.

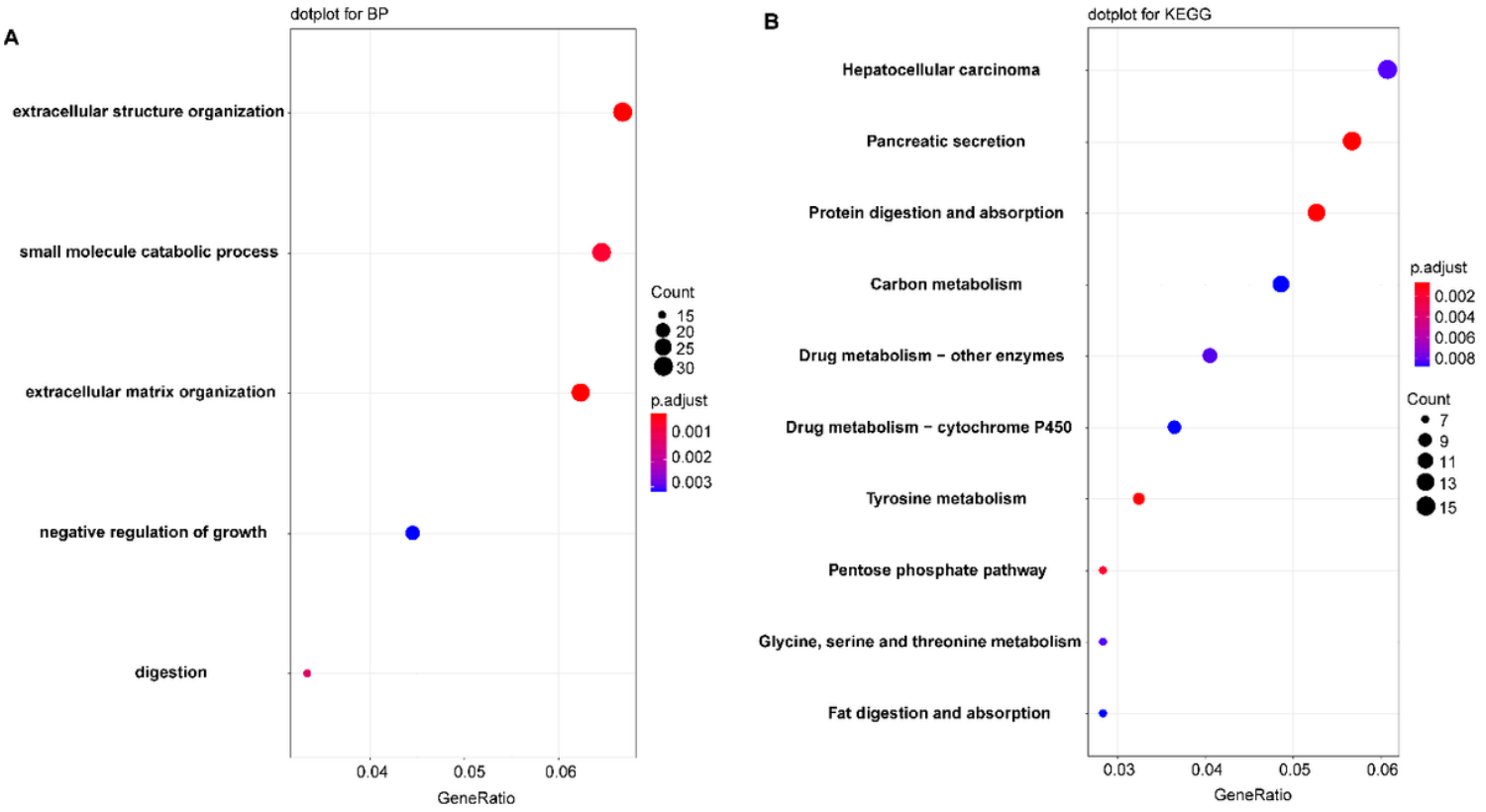

Figure 2

Functional analysis. (A) Gene Ontology analysis for genes in yellow module. (B) KEGG pathway enrichment analysis for genes in yellow module. The x-axis shows the -log10 (P- value) of each term and the y-axis shows the GO and KEGG pathway terms.

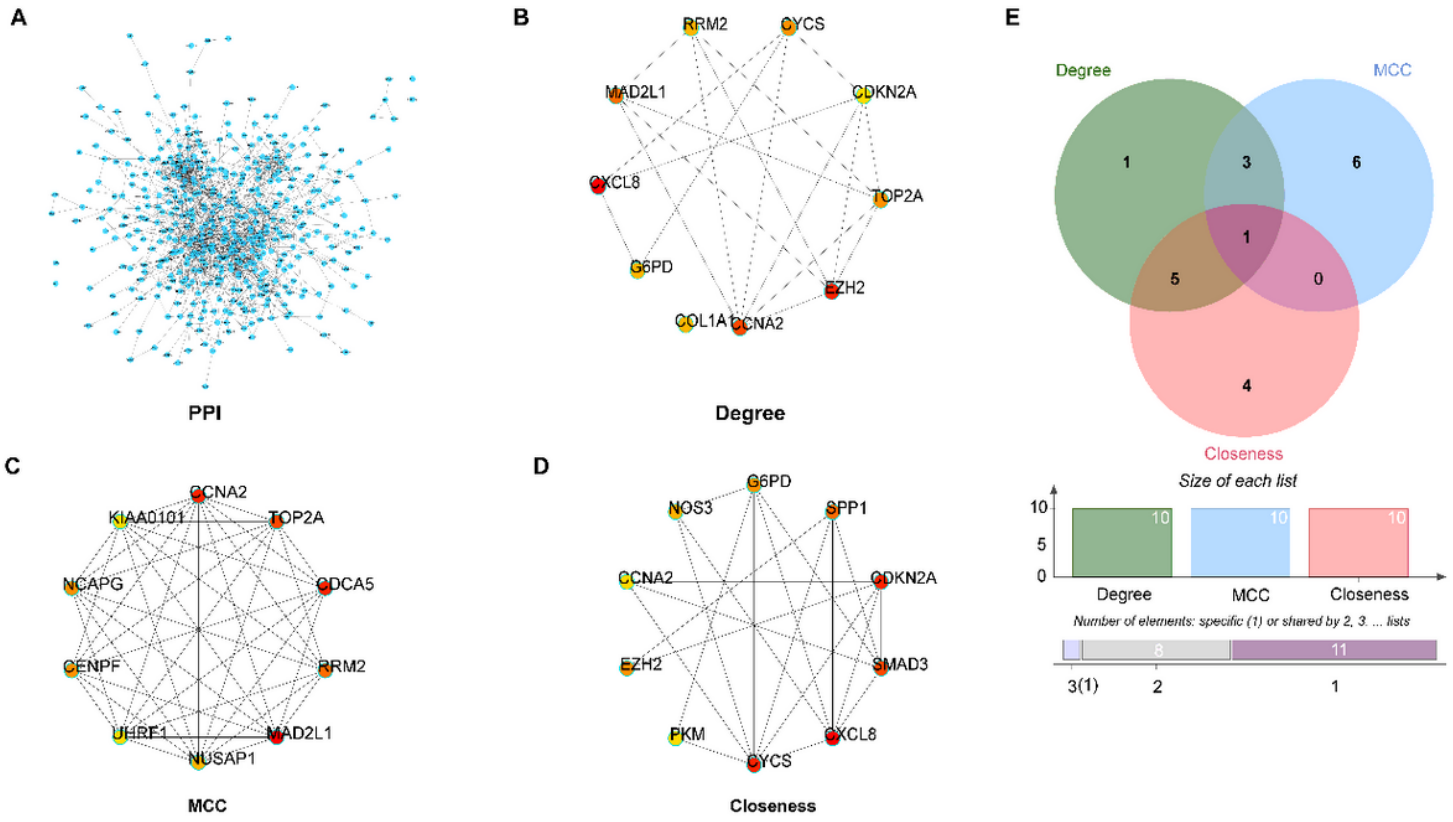

Figure 3

Hub genes detection and protein-protein network (PPI). (A) PPI network of differentially expressed genes. (B) (C) (D) The top 10 hub genes were depicted in a network using the CytoHubba plugin. (E) Selection of hub genes in PPI network and co-expression network.

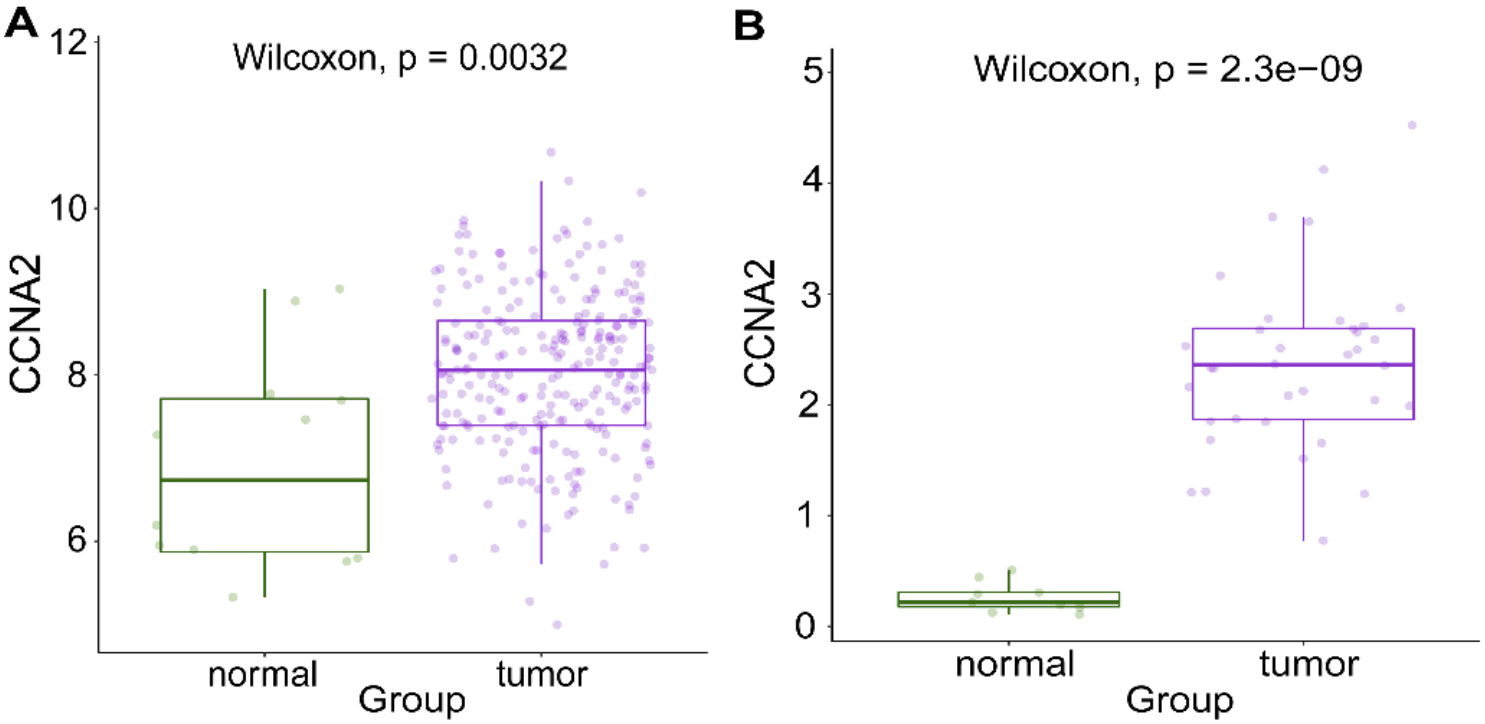

Figure 4

Comparison of CCNA2 expression between CCA and adjacent noncancerous tissue in (A) GEO database and (B)TCGA database.

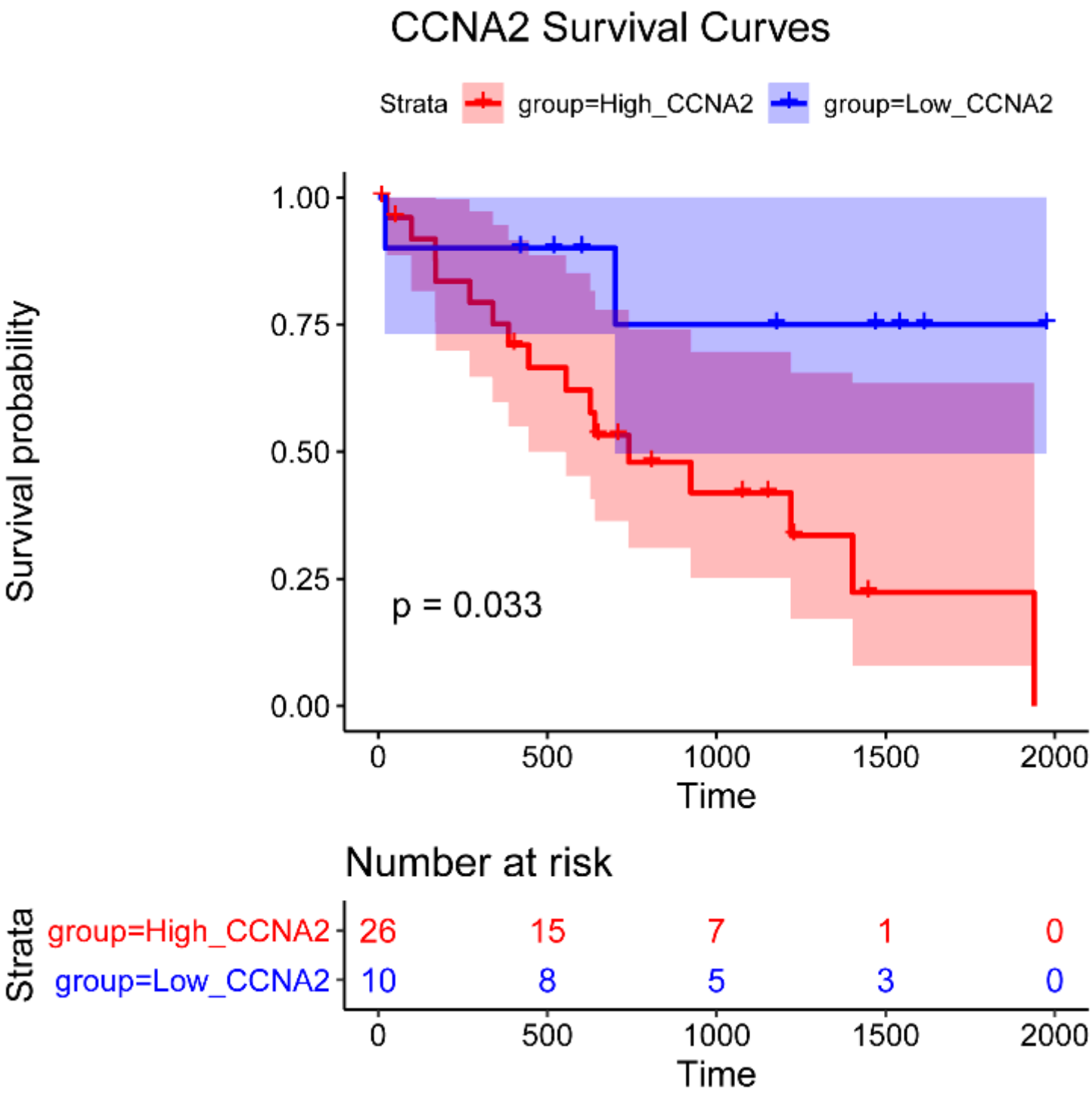

Figure 5

Kaplan–Meier curve of the relationship between CCNA2 expression and the prognosis of CCA patients based on TCGA database.

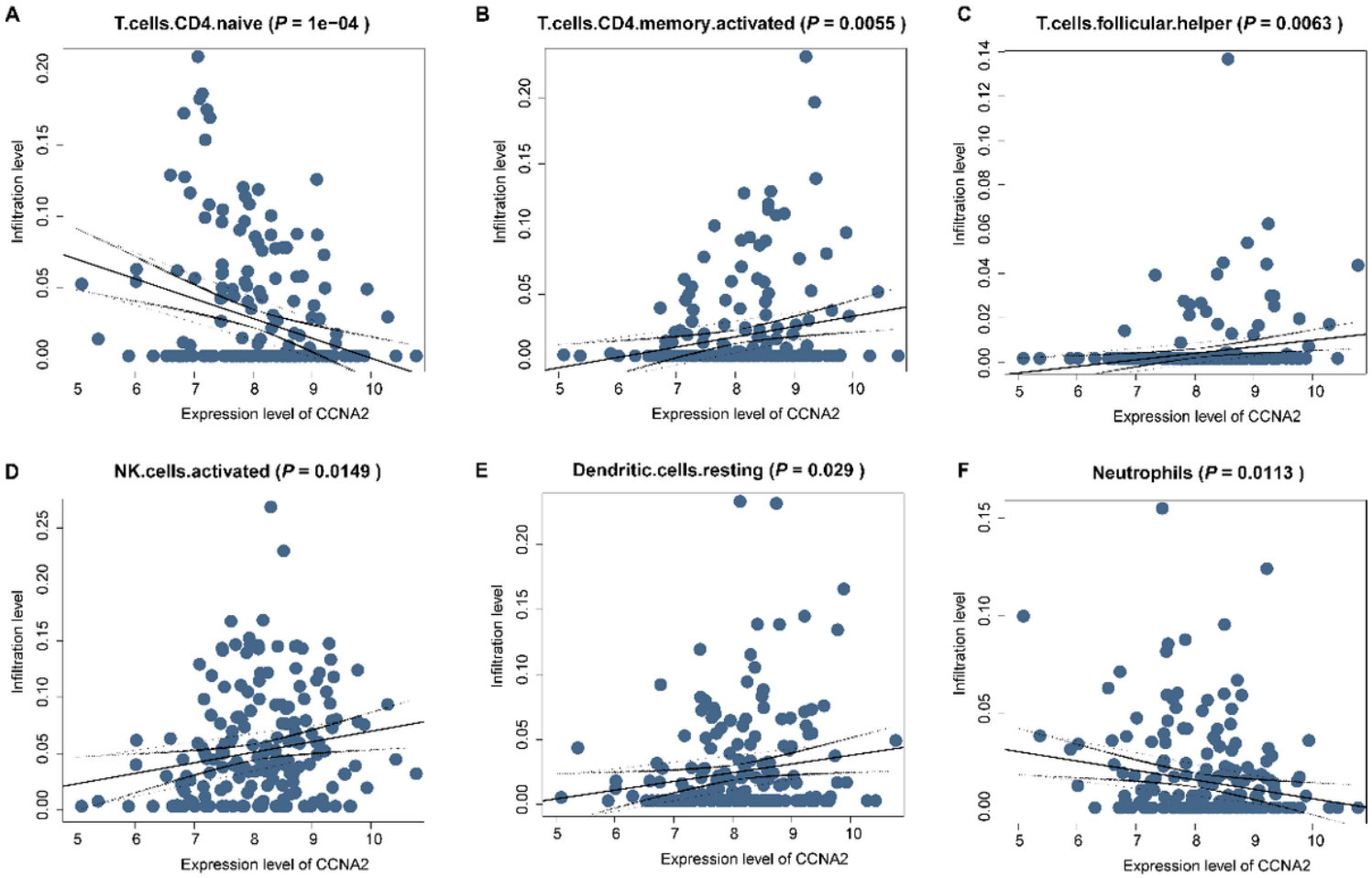

Figure 6

The co-expression patterns among fractions of immune cell. (A)T cells CD4 naive and CCNA2; (B)T cells CD4 memory and CCNA2; (C) T cells follicular and CCNA2; (D)NK cells and CCNA2; (E)Dendritic cells and CCNA2; (F) Neutrophils and CCNA2

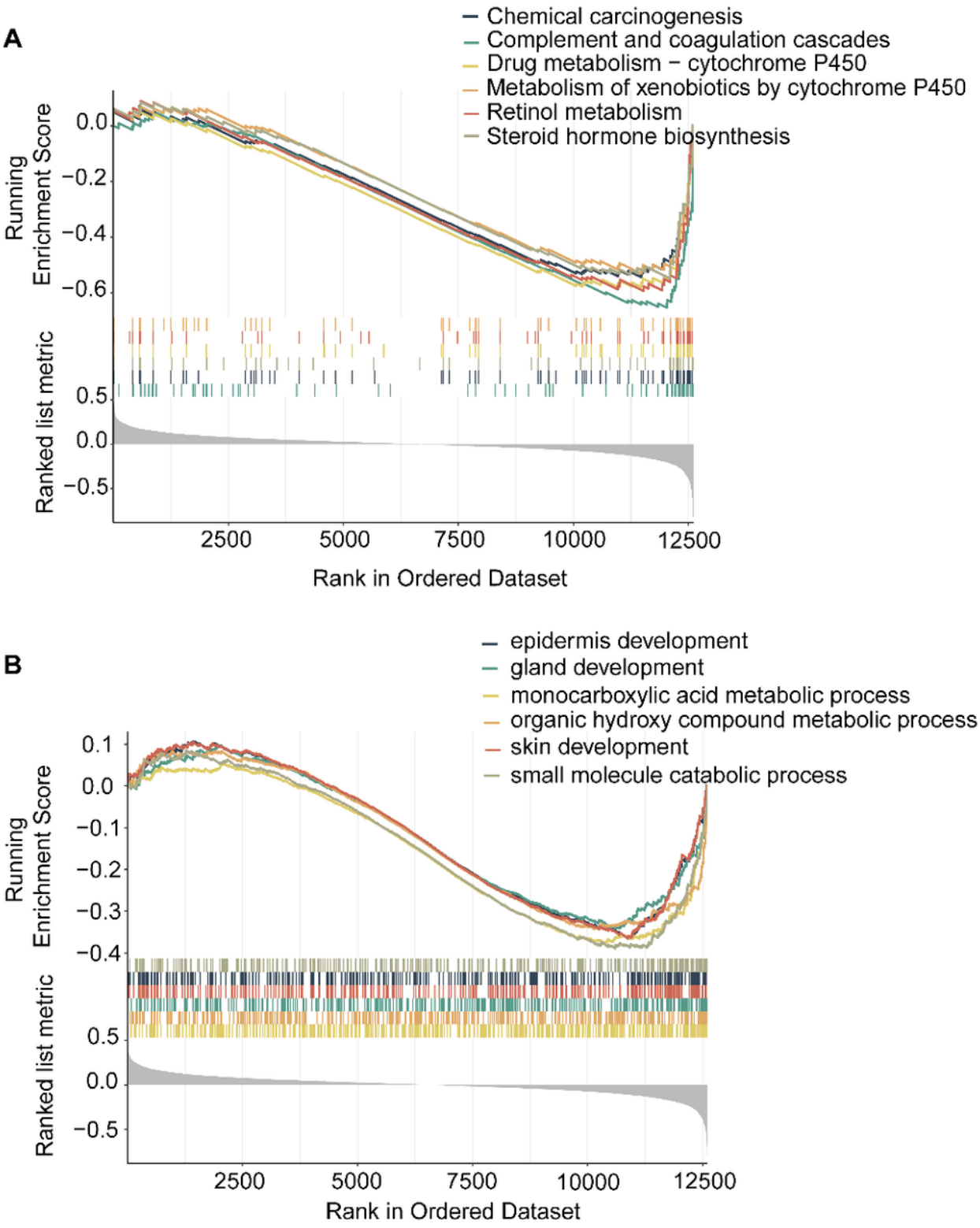

Figure 7

A merged enrichment plot from gene set enrichment analysis including enrichment score and gene sets.