

Analysis of differential gene expression between right-sided and left-sided colon cancer by bioinformatics analysis

Qiangwei Chi

Wenzhou Medical University First Affiliated Hospital

Shizuan Chen

Wenzhou Medical University First Affiliated Hospital

Shaotang Li (✉ lishaotang@wmu.edu.cn)

Wenzhou Medical University First Affiliated Hospital

Research

Keywords: bioinformatics analysis, colon cancer, tumor locations, biomarkers, immunotherapy

Posted Date: April 23rd, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-23894/v1>

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

[Read Full License](#)

Abstract

Background

Colon cancer is a common tumor of the digestive tract worldwide. Recent researches have revealed that colon cancer exhibits distinct differences in clinical and biological characteristics depending on the location of the tumor. However, the underlying genetic and molecular mechanism of the differences between right-sided colon cancer (RCC) and left-sided colon cancer (LCC) are not fully understood. This study aimed to identify molecular potential biomarkers and therapeutic targets for precise treatment of right-sided and left-sided colon cancer using bioinformatics analysis.

Methods

The gene microarray profile, named GSE44076, from the Gene Expression Omnibus (GEO) public database was downloaded and processed to then select differentially expressed genes (DEGs) on the base of two sample groups of RCC and LCC. Also, gene ontology (GO) analysis, Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis, protein–protein interaction (PPI) network construction, module analysis, validation of hub genes, and survival analysis.

Results

Finally, we obtained 2259 DEGs between RCC and LCC, 1300 of which were upregulated in RCC and 945 of which were upregulated in LCC. The results of GO and KEGG analysis of the DEGs indicated that the biological functions of DEGs in RCC and LCC were significantly different. *CTLA4*, *IL10*, *IL2RB*, *IFNG*, *NCAM1*, *EGFR*, *MYC*, *SRC*, *CUL3*, and *NCBP2* were identified from the PPI networks as the hub genes of RCC and LCC. Among the hub genes, the log-rank tests for overall survival (OS) and disease free survival (DFS) were applied. Moreover, all hub genes, except *CUL3*, had differential expression levels of miRNA between tumor group and normal group.

Conclusion

These hub genes and pathways identified based on bioinformatics analysis might conduce to explain the differences between RCC and LCC, and most of the hub genes were specific to the malignant tissues. Notably, these hub genes, especially the genes associated with immunotherapy such as *CTLA4*, might be potential specific targets or prognostic markers for precise treatment of colon cancer.

Background

Colon cancer is a common malignant tumor of the digestive system globally, with increasing incidence and mortality [1]. Colon cancer can be approximately divided into right-sided colon cancer (RCC) and left-

sided colon cancer (LCC) according to the anatomical site of the tumor. To date, a growing body of researches have demonstrated that LCC and RCC differ distinctly in terms of the epidemiology, clinical feature, drug sensitivity, treatment strategy, as well as prognosis, and some researchers deemed that RCC and LCC were supposed to be even regarded as two distinct diseases [2, 3]. Also, recent studies have shown that there were underlying molecular differences between RCC and LCC. High microsatellite instability (MSI-H) or mismatch repair deficient (dMMR) tumors are more commonly located in RCC, which might act a pivotal part in the chemotherapy and outcome of RCC [4], whereas chromosome instability (CIN) and p53 mutations were characteristic of LCC [5]. However, the deeper molecular and genetic mechanism of the RCC and LCC was not fully understood. Thus, in-depth studies into colon cancers with different tumor locations were warranted for improving the precise treatment of colon cancers.

Microarray, as a high-throughput sequencing method, has been widely used to obtain information about cancer gene expression profiles in life science [6]. Nowadays, the amount of sequence data has been growing at an unprecedented magnitude worldwide. Unfortunately, vast amounts of these sequence data were just stored in various online databases after program completion. Hence, re-analyzing these data from sequencing technology might constitute a cheap and useful way to understand tumorigenesis and cancer development.

In this study, we aimed to reveal the molecular genetic mechanisms responsible for biological phenotypic differences in RCC and LCC by bioinformatics analysis. First, we downloaded the gene microarray profile (GSE44076) and used the Limma R package to identify the differentially expressed genes (DEGs) based on two sample groups of RCC and LCC. After analyzing the DEGs by a series of bioinformatics methods, we identified several hub genes and pathways from the PPI networks of RCC and LCC. Our research will provide some new insights into colon cancer with different anatomic sites.

Methods

Microarray data

The gene microarray profile (GSE44076) was downloaded from the GEO database of the National Center for Biotechnology Information (NCBI) [7]. The dataset contains 98 colon cancer samples, including 60 LCC and 38 RCC patients. All samples were from early-stage (IIA-B) colon cancer patients diagnosed by pathologic evidence. Meanwhile, clinical data of these patients in the microarray profile was collected and evaluated in this study.

Data processing and screen for DEGs

R 3.6.0 software was utilized to normalize the dataset through the `normalizeBetweenArrays` function of the Limma R package [8], which can eliminate batch effects of the samples. Meanwhile, gene expression data of all samples were subjected to log₂ transformation. Each probe ID of data was then converted to the corresponding official gene symbols by annotation of GPL570 platform; if multiple probe IDs mapped

to the same gene symbol, then choose the max expression value representing that gene's expression. In order to obtain sufficient numbers of DEGs in this paper, we used a relaxed cutoff criterion (a corrected $P < 0.05$ without the limitation of absolute fold change value) for statistical analyses. Eventually, the Limma R package was applied to identify DEGs by comparing expression value between LCC and RCC.

GO and KEGG analysis

The Database for Annotation, Visualization and Integrated Discovery (DAVID version 6.8, <https://david.ncifcrf.gov/>) is a publicly available database for gene functional annotation and pathway enrichment analysis. In this study, we mapped the DEGs above into the DAVID to carry out the KEGG pathway analyses and GO analyses, respectively. Any adjusted P-value less than 0.05 was considered as statistically significant. Then, visualization for results of GO and KEGG analysis was performed in R software.

PPI network and modules analysis

String (<https://string-db.org/>) is an online software that can analyze the interaction relationships between the DEGs. Based on this software, PPI networks were constructed for LCC and RCC, respectively. Subsequently, the PPI networks were visualized by Cytoscape 3.6.1. According to the degrees, the top 5 genes were screened as hub genes of RCC and LCC. In addition, modules in the PPI networks were explored using the plug-in Molecular Complex Detection (MCODE) of Cytoscape. The KEGG pathway enrichment analysis of the DEGs in the modules was also carried out using DAVID tool.

Validation of the hub genes

In order to validate that these hub genes were not from potential baseline differences between right and left colon tissues, we verified the miRNA expression levels of all the hub genes between tumor tissues and non-tumor tissues.

Survival analysis for hub genes

The Gene Expression Profiling Interactive Analysis (GEPIA, <http://gepia.cancer-pku.cn>) is an online service that provides overall survival (OS) or disease free survival (DFS, also called relapse-free survival and RFS) analysis of single or multiple genes. The survival analysis for hub genes was performed using this tool. Select the quartile in gene expression for splitting the high-expression and low-expression cohorts. The 95% confidence intervals were calculated and presented as dotted lines on the plot. The log-rank P-values less than 0.05 were considered to be significant.

Results

Baseline characteristics

In this study, we collected clinical data of 38 RCC and 60 LCC samples in GSE44076. The baseline characteristics of the samples were shown in Table 1. There were no significant differences in any of the baseline characteristics between the two groups, which guarantees a good comparability.

Data processing and identification of DEGs

First, the colon cancer gene microarray data was normalized by `normalizeBetweenArrays` function (Fig. 1). The DEGs between RCC and LCC were identified using the Limma R package. A total of 2259 DEGs were obtained from the GSE44076 dataset, 1300 of which were upregulated in RCC and 945 of which were upregulated in LCC. The top 100 DEGs in RCC and LCC were presented in Table 2.

GO and KEGG analysis

GO analysis is a commonly used method for the functional annotation of genes and gene products. Based on the DAVID database, GO analysis covers three categories to describe biological functions of DEGs: biological process (BP), molecular function (MF) and cell component (CC). The DEGs upregulated in RCC were mainly enriched in immune response, external side of plasma membrane, and transferase activity (Fig. 2a), whereas DEGs upregulated in LCC were mainly involved in protein ubiquitination, cytosol, and protein binding (Fig. 2b).

The KEGG database is one of the most commonly used bioinformatics databases. In this study, we used KEGG analysis to explore cell functions and molecular functions (MF) of the DEGs at a deeper level. The DEGs upregulated in RCC were mainly enriched in antigen processing and presentation, T cell receptor signaling pathway, cytokine-cytokine receptor interaction (Fig. 3a), while the DEGs upregulated in LCC were mainly involved in endocytosis and TGF-beta signaling pathway (Fig. 3b).

PPI network construction and modules analyze

The STRING online database was utilized to construct PPI networks of RCC and LCC respectively, and Fig. 4 illustrated that all nodes degree distribution of PPI networks. Then, the top 5 hub genes in PPI networks were screened through ranking node degrees. The hub genes identified from the RCC network were *CTLA4*, *IL10*, *IL2RB*, *IFNG*, and *NCAM1*, whereas the hub genes in LCC were *EGFR*, *MYC*, *SRC*, *CUL3*, and *NCBP2*, as shown in Table 3. Moreover, the top modules of the PPI networks for RCC and LCC were selected using MCODE and the KEGG analysis of the genes in two modules exhibited that these genes were mainly enriched in T cell receptor signaling pathway, measles, and ubiquitin-mediated proteolysis (Table 4).

The mRNA expression levels of the hub genes

All these hub genes, except *CUL3*, had significant differential expression levels of miRNA between tumor group and normal group from GSE44076 data, and Fig. 5 showed that the gene expression levels of *CTLA4*, *IL10*, *IFNG*, *MYC*, *SRC*, and *NCBP2* were up-regulated in colon cancer samples compared to normal tissues, whereas *IL2RB*, *NCAM1*, and *EGFR* were decreased in cancer patients.

Survival analysis

We used the GEPIA tool to perform OS and DFS analysis for these hub genes. The results of survival analysis showed that *EGFR* and *IFNG* overexpression were found to be connected with poor DFS in colon cancer patients, and lower expression of *NCBP2* was associated with poor OS (Fig. 6).

Discussion

Taking into account that advanced protocols of chemotherapy and targeted therapy, the treatment of colon tumors should not only be decided by clinical stage and pathological grade of the tumor but the molecular mechanism of tumorigenesis, some of which were associated with the location of the tumor. The tumor location could contribute to the differences in the pathological type, clinical manifestations, prognosis and chemoresistance in the colon cancer. It is, therefore, essential to regard the tumor location as an important factor in the management of colon cancer, which might offer personalized care and accurate treatment for cancer patients.

Presently, several studies have revealed that there are differences in colon cancers with different tumor locations at the aspect of clinical features, but the underlying molecular genetic mechanism has not yet been elucidated [9–12]. In this paper, we explored potential gene expression differences between RCC and LCC using bioinformatics analysis methods.

The gene chip GSE44076 we selected was composed of a series homogeneous of early-stage (A-B) colon cancer samples. A total of 2259 DEGs, 1300 of which were upregulated in RCC and 945 of which were upregulated in LCC, were obtained. To prove the functional heterogeneity of DEGs between RCC and LCC, we then performed GO annotation and KEGG pathway analysis on these DEGs to explore higher-level functional differences. DEGs upregulated in RCC were mainly enriched in immune response, external side of plasma membrane, and transferase activity, whereas DEGs upregulated in LCC were mainly involved in protein ubiquitination, cytosol, and protein binding. The KEGG pathways of upregulated DEGs in RCC were mainly enriched in antigen processing and presentation, T cell receptor signaling pathway, cytokine-cytokine receptor interaction, while the DEGs upregulated in LCC were mainly involved in endocytosis and Hippo signaling pathway. Among these pathways, Hippo signaling pathway had been widely studied in the tumor field, which might affect the metastasis, drug resistance, and recurrence of colorectal cancer [13]. Moroish et al. had found that loss of the Hippo pathway kinases LATS1/2 (large tumor suppressor 1 and 2) in tumor cells inhibits tumor growth by enhancing anti-tumor immune responses, suggesting that the pathway might be a potential therapeutic target for LCC [14]. Besides, we noticed that the enrichment pathways of DEGs upregulated in RCC were involved in immune responses and cytokine interaction, indicating that immune responses are more active in RCC. Here, GO and KEGG analysis revealed that there were different cellular functional levels and signaling pathways in the two groups of DEGs, which could help us to re-recognize the differences in colon cancer with different anatomic sites.

Then, we constructed PPI networks with the two groups of DEGs and screened the following 10 hub genes: *CTLA4*, *IL10*, *IL2RB*, *IFNG*, *NCAM1*, *EGFR*, *MYC*, *SRC*, *CUL3*, and *NCBP2*. Then, we found that there

were significant expression differences in these hub genes between colon cancer tissues and normal tissues which provided evidence to some extent that the hub genes were specific to the malignant tissues and were not from potential baseline differences between right and left colon tissues. We also performed a prognosis analysis of hub genes using the GEPIA online software. Survival analysis indicated that three of the hub genes, *EGFR*, *NCBP2*, and *IFNG*, were correlated with the OS or DFS of colon cancer patients. To a certain extent, these findings might explain the survival difference in RCC and LCC.

CTLA4 is the receptor that negatively regulates T-cell activation by suppressing glucose metabolism and Akt activity [15]. Based on the concept of immune checkpoints, *CTLA4* has become one of the most popular therapeutic targets for tumor immunotherapy in recent years, and CTLA4 inhibitor Ipilimumab is already approved by FDA for the treatment of MSI-H/dMMR metastatic colorectal cancer [16]. Considering that MSI-H/dMMR is commonly located in RCC and *CTLA4* gene was identified as the hub gene of RCC, we speculate that CTLA4 inhibitor might be more beneficial for the treatment of RCC. IL10 is a cytokine produced mainly by macrophages, regulatory T cells and epithelial cells [17]. IL10 has been considered to have immunosuppressive and tumor-promoting potentials in the past [18]. However, it was recently reported that IL-10 was able to promote tumor immunity and developed as an anti-tumor drug for the treatment of cancer patients [19]. Naing et al. have found that PEGylated IL-10 can induce CD8 + T cell immunity in cancer patients and promote the expansion of underrepresented T cell clones, which would improve the therapeutic efficacy of anti-PD-1 [20]. *IL2RB* gene encodes the IL2 receptor beta chain, a component of the IL2 receptor, which can combine with IL2 to regulate T cell activation and proliferation [21]. Kagoya et al. constructed a novel chimeric antigen receptor T cell (CAR-T) using a cytoplasmic domain of IL2 receptor β , called 28- Δ IL2RB-z (YXXQ) CAR-T cells, which suggested superior persistence and antitumor effects in both liquid and solid tumor models compared with traditional CAR-T cells [22]. Besides, according to the KEGG pathways analysis, the DEGs of RCC in the selected module, including *CTLA4*, *IL10*, and *IFNG*, were enriched in T cell receptor signaling pathway, suggesting that T cell receptor signaling pathway might be a therapeutic target to suppress colon cancer. Overall, these findings may be important in the treatment of colon cancer, supporting that tumor immunotherapy could act as a potential treatment strategy for RCC, such as anti-CTLA4, anti-PD-1, and CAR-T.

EGFR, as an important regulatory factor in tumor cells, was demonstrated to facilitate tumor cell progression by activating Ras/MAPK and PI3K/Akt signaling pathway [23]. To date, anti-EGFR therapy has become an important part of anti-tumor therapy for colon cancer. Recent studies have shown that patients with distal colon cancers have a better prognosis respond to anti-EGFR therapy [24, 25]. Moreover, the results of our study demonstrated that a significantly higher level of *EGFR* expression was found in LCC sample, indicating that the better prognosis with anti-EGFR therapy in LCC patients may be related to the expression of *EGFR* was higher in LCC. *SRC*, the first proto-oncogene found, encodes a cytoplasmic tyrosine kinase that belongs to Src family kinases, which were found to play a key role in the growth, differentiation, proliferation, invasion, and metastasis of tumor cells [26]. A cohort study demonstrated that higher mRNA expression of *Src* was related to poor prognosis in early-stage colon cancer and had a stronger negative impact on outcome in LCC and stage II disease [27]. Several studies have found a more favorable survival adjusted for tumor stage in LCC [28–30], whereas Weiss et al.

indicated that RCC in stage I-II disease demonstrated a better overall survival compared to LCC [31], which might be related to the expression of *SRC* was higher in early-stage LCC. Besides, it was reported that *SRC* overexpression was associated with colon cancer resistance to oxaliplatin and SRC inhibitors can restore sensitivity to oxaliplatin in tumors with high levels of phospho-Src [32]. These findings supported that *SRC* might be a novel biological marker for prognosis analysis and chemotherapy effects of colon cancer patients, particularly LCC.

Conclusion

In summary, this study was aimed to identify DEGs between RCC and LCC and construct corresponding functional networks and signal pathways through integrated bioinformatics analysis to provide some new insights into the molecular mechanisms of colon cancers with different tumor locations. In addition, we identified some hub genes and signal pathways that could be potential biomarkers or therapeutic targets for LCC and RCC. However, the results of this study require further molecular biological experiments to verify.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The data that support the findings of this study are available from GEO database.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by Wenzhou science and technology bureau project of China (y20180085).

Authors' contributions

QW conceived the study, analysed the data and wrote this manuscript. ST designed the study, performed the research. SZ performed the part of research and acquired the data. All authors read and approved the final manuscript.

Acknowledgements.

Not applicable.

References

1. Labianca R, Beretta GD, Kildani B, Milesi L, Merlin F, Mosconi S, Pessi MA, Prochilo T, Quadri A, Gatta G. Colon cancer. *Crit Rev Oncol Hematol*. 2010;74:106-33.
2. Loupakis F, Yang D, Yau L, Feng S, Cremolini C, Zhang W, Maus MK, Antoniotti C, Langer C, Scherer SJ. Primary tumor location as a prognostic factor in metastatic colorectal cancer. *JNCI: Journal of the National Cancer Institute*. 2015;107.
3. Hussain M, Waqas O, Hassan U, Loya A, Akhtar N, Mushtaq S, Yusuf MA, Syed AA. Right-sided and left-sided colon cancers are two distinct disease entities: an analysis of 200 cases in Pakistan. *Asian Pac J Cancer Prev*. 2016;17:2545-8.
4. Bertagnolli MM, Niedzwiecki D, Compton CC, Hahn HP, Hall M, Damas B, Jewell SD, Mayer RJ, Goldberg RM, Saltz LB. Microsatellite instability predicts improved response to adjuvant therapy with irinotecan, fluorouracil, and leucovorin in stage III colon cancer: Cancer and Leukemia Group B Protocol 89803. *J Clin Oncol*. 2009;27:1814.
5. Pritchard CC, Grady WM. Colorectal cancer molecular biology moves into clinical practice. *Gut*. 2011;60:116-29.
6. Lu X, Zhang X. The effect of GeneChip gene definitions on the microarray study of cancers. *Bioessays*. 2006;28:739-46.
7. Sanz-Pamplona R, Berenguer A, Cordero D, Molleví DG, Crous-Bou M, Sole X, Paré-Brunet L, Guino E, Salazar R, Santos C. Aberrant gene expression in mucosa adjacent to tumor reveals a molecular crosstalk in colon cancer. *Mol Cancer*. 2014;13:46.
8. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res*. 2015;43:e47-e.
9. Beart RW, Melton LJ, Maruta M, Dockerty MB, Frydenberg HB, O'Fallon WM. Trends in right and left-sided colon cancer. *Dis Colon Rectum*. 1983;26:393-8.
10. Nawa T, Kato J, Kawamoto H, Okada H, Yamamoto H, Kohno H, Endo H, Shiratori Y. Differences between right-and left-sided colon cancer in patient characteristics, cancer morphology and histology. *J Gastroenterol Hepatol*. 2008;23:418-23.
11. Gao P, Song Yx, Xu Yy, Sun Z, Sun Jx, Xu Hm, Wang Zn. Does the prognosis of colorectal mucinous carcinoma depend upon the primary tumour site? Results from two independent databases. *Histopathology*. 2013;63:603-15.
12. Cremolini C, Antoniotti C, Lonardi S, Bergamo F, Cortesi E, Tomasello G, Moretto R, Ronzoni M, Racca P, Loupakis F. Primary tumor sidedness and benefit from FOLFOXIRI plus bevacizumab as initial therapy for metastatic colorectal cancer. Retrospective analysis of the TRIBE trial by GONO. *Ann Oncol*. 2018;29:1528-34.

13. Poma AM, Torregrossa L, Bruno R, Basolo F, Fontanini G. Hippo pathway affects survival of cancer patients: extensive analysis of TCGA data and review of literature. *Sci Rep.* 2018;8:1-12.
14. Moroishi T, Hayashi T, Pan W-W, Fujita Y, Holt MV, Qin J, Carson DA, Guan K-L. The Hippo pathway kinases LATS1/2 suppress cancer immunity. *Cell.* 2016;167:1525-39. e17.
15. Parry RV, Chemnitz JM, Frauwirth KA, Lanfranco AR, Braunstein I, Kobayashi SV, Linsley PS, Thompson CB, Riley JL. CTLA-4 and PD-1 receptors inhibit T-cell activation by distinct mechanisms. *Mol Cell Biol.* 2005;25:9543-53.
16. Overman MJ, Lonardi S, Wong KYM, Lenz H-J, Gelsomino F, Aglietta M, Morse MA, Van Cutsem E, McDermott R, Hill A. Durable clinical benefit with nivolumab plus ipilimumab in DNA mismatch repair-deficient/microsatellite instability-high metastatic colorectal cancer. 2018.
17. Ouyang W, Rutz S, Crellin NK, Valdez PA, Hymowitz SG. Regulation and functions of the IL-10 family of cytokines in inflammation and disease. *Annu Rev Immunol.* 2011;29:71-109.
18. Sharma S, Stolina M, Lin Y, Gardner B, Miller PW, Kronenberg M, Dubinett SM. T cell-derived IL-10 promotes lung cancer growth by suppressing both T cell and APC function. *The Journal of Immunology.* 1999;163:5020-8.
19. Teng MW, Darcy PK, Smyth MJ. Stable IL-10: a new therapeutic that promotes tumor immunity. *Cancer Cell.* 2011;20:691-3.
20. Naing A, Infante JR, Papadopoulos KP, Chan IH, Shen C, Ratti NP, Rojo B, Autio KA, Wong DJ, Patel MR. PEGylated IL-10 (Pegilodecakin) induces systemic immune activation, CD8+ T cell invigoration and polyclonal T cell expansion in cancer patients. *Cancer Cell.* 2018;34:775-91. e3.
21. Lin J-X, Li P, Liu D, Jin HT, He J, Rasheed MAU, Rochman Y, Wang L, Cui K, Liu C. Critical Role of STAT5 transcription factor tetramerization for cytokine responses and normal immune function. *Immunity.* 2012;36:586-99.
22. Kagoya Y, Tanaka S, Guo T, Anczurowski M, Wang C-H, Saso K, Butler MO, Minden MD, Hirano N. A novel chimeric antigen receptor containing a JAK-STAT signaling domain mediates superior antitumor effects. *Nat Med.* 2018;24:352.
23. Baselga J, Arteaga CL. Critical update and emerging trends in epidermal growth factor receptor targeting in cancer. *J Clin Oncol.* 2005;23:2445-59.
24. Brule S, Jonker D, Karapetis C, O'Callaghan CJ, Moore M, Wong R, Tebbutt NC, Underhill C, Yip D, Zalcberg JR. Location of colon cancer (right-sided versus left-sided) as a prognostic factor and a predictor of benefit from cetuximab in NCIC CO. 17. *Eur J Cancer.* 2015;51:1405-14.
25. Wang F, Bai L, Liu T-S, Yu Y-Y, He M-M, Liu K-Y, Luo H-Y, Zhang D-S, Jin Y, Wang F-H. Right-and left-sided colorectal cancers respond differently to cetuximab. *Chin J Cancer.* 2015;34:24.
26. Summy JM, Gallick GE. Src family kinases in tumor progression and metastasis. *Cancer Metastasis Rev.* 2003;22:337-58.
27. Martínez-Pérez J, Lopez-Calderero I, Saez C, Benavent M, Limon ML, Gonzalez-Exposito R, Soldevilla B, Riesco-Martínez MC, Salamanca J, Carnero A. Prognostic relevance of Src activation in stage II-III colon cancer. *Hum Pathol.* 2017;67:119-25.

28. Suttie S, Shaikh I, Mullen R, Amin A, Daniel T, Yalamarthy S. Outcome of right-and left-sided colonic and rectal cancer following surgical resection. *Colorectal Dis.* 2011;13:884-9.
29. Wray CM, Ziogas A, Hinojosa MW, Le H, Stamos MJ, Zell JA. Tumor subsite location within the colon is prognostic for survival after colon cancer diagnosis. *Dis Colon Rectum.* 2009;52:1359-66.
30. Derwinger K, Gustavsson B. Variations in demography and prognosis by colon cancer location. *Anticancer Res.* 2011;31:2347-50.
31. Weiss JM, Pfau PR, O'Connor ES, King J, LoConte N, Kennedy G, Smith MA. Mortality by stage for right-versus left-sided colon cancer: analysis of surveillance, epidemiology, and end results–Medicare data. *J Clin Oncol.* 2011;29:4401.
32. Perez M, Lucena-Cacace A, Marín-Gómez LM, Padillo-Ruiz J, Robles-Frias MJ, Saez C, Garcia-Carbonero R, Carnero A. Dasatinib, a Src inhibitor, sensitizes liver metastatic colorectal carcinoma to oxaliplatin in tumors with high levels of phospho-Src. *Oncotarget.* 2016;7:33111.

Tables

Table 1. Clinical characteristics of samples in RCC and LCC.

Clinical characteristics	RCC	LCC	P-value
Samples	38	60	-
Mean age(years)	70.5±9.0	70.6±9.1	0.972
Sex			
Male	27 [71.1%]	44 [73.4%]	0.989
Female	11 [28.9%]	16 [26.7%]	
Clinical stages			
A	36 [94.7%]	54 [90%]	0.479
B	2 [5.3%]	6 [10%]	

Table 2. The top 100 DEGs from GSE44076 ranked by adjusted P-value.

DEGs	Gene symbol
Upregulated in RCC	<i>HOXB6, GABRP, ZNF672, GF11, LCK, APOL4, HOXB5, DMRTA2, HEXDC, ISOC2, GPR18, OSR2, PATL2, CD3G, PNPLA3, TMEM229B, HSD3B2, NLRC5, KSR2, GSG1, L1TD1, BARX2, TMPRSS3, GSDMB, SLC28A3, BCL2, PLEKHG7, ASRGL1, RASGRP1, CYP4X1, C8orf86, LYPD5, PTGER2, WIBG, GPR171, HOXC6, MS4A8B, ELOVL7, TRNP1, BAK1, RAB26, CCDC68, VPS37B, ETNK1, RILPL2, COQ10A, GBA3, TRAF3IP2, C10orf81, CD5, CYTIP, KLK3, ZNF407, PML, PVRL2, ABCG5, CA12, HOXB8, HTR3E, CLDN2, CDKL1, RNF138, TMEM61, KCNJ2, CASP1, L3MBTL3, MYEOV, WDR54, AVIL, DRD5, BMX, MOSC1, FOXA1, FOXI1, C1orf226, MB, SMAP2, FKRP, TMC8, CD2, RARRES3, ESRRG, DUSP4, DLG3, TRERF1, DHRS4, C19orf54, METTL12, NEURL, TTC39C, ARX, VPS13A, FOXD4, FGD4, TOX, ATP2B3, C14orf159, HIP1R, CXCR6, RNF183, FAR2</i>
Upregulated in LCC	<i>COPS8, MFF, SPAG16, NFE2L2, TP53RK, PFDN4, TEAD3, ATP6V1C1, MFF, MAPRE1, HDGFRP3, FITM2, ELAVL2, APLF, CEBPB, ZNF347, PRAC, IRAK2, Schr2q13, SERINC3, WASF3, GTF2E1, SH3GL1, LOC647979, Septin10, C20orf177, WDFY1, ZFP28, NEDD4, GPX4, AK3L1, DBNDD2, KIAA0196, DYNLRB1, HEBP2, PPP1R3D, CXCR7, SUPT7L, ITCH, C20orf24, TM9SF4, STAU1, NCOA3, C6orf1, CDYL, UBE2V1, NCBP2, IL28RA, THRB, PLCG1, AMOTL2, RAB32, TTPAL, FOXO1, C7orf25, FAM101A, MOCS3, AGFG1, ATP5E, ZMYM5, FBXL12, SDC4, FDX1L, TNNC1, RALB, CRIPT, TCFL5, NSMCE2, ROMO1, ADAM17, SERINC1, SKIL, C20orf111, ZSCAN18, PPT2, PDP1, TTC35, ATP6V1H, RGL2, TBC1D23, ITGB5, KBTBD2, DNNTIP1, PPP1CB, LOC645405, EREG, ASAP1, ZNF285, RANBP9, RB1CC1, SYS1, CCDC90A, FER1L4, APLF, RALGAPB, KRTAP33, ATG3, RPRD1B, STX16, E2F6, CTSL2</i>

Table 3. Top 10 hub genes with higher degree of connectivity in RCC and LCC.

Gene symbol	Gene description	Degree	Adj.P-value
RCC			
CTLA4	Cytotoxic T-lymphocyte associated protein 4	60	0.011
IL10	Interleukin 10	56	0.012
IL2RB	Interleukin 2 receptor subunit beta	49	0.002
IFNG	Interferon gamma	49	0.015
NCAM1	Neural cell adhesion molecule 1	45	0.011
LCC			
EGFR	Epidermal growth factor receptor	81	0.006
MYC	MYC proto-oncogene	68	0.010
SRC	SRC proto-oncogene	64	0.010
CUL3	Cullin 3	39	0.012
NCBP2	Nuclear cap binding Protein subunit 2	27	0.000

Table 4. The KEGG analysis of top modules from the PPI network in RCC and LCC.

Module	Pathway	Count	P.value	Genes
RCC_Module	T cell receptor signaling pathway	8	1.25E-09	<i>ITK, CD3G, CD3D, FYN, IFNG, ZAP70, CTLA4, IL10</i>
	Measles	6	8.17E-06	<i>IL2RB, CD3G, IL2RA, CD3D, FYN, IFNG</i>
LCC_Module	Ubiquitin mediated proteolysis	6	1.96E-08	<i>CUL3, NEDD4, UBE2W, KEAP1, ITCH, UBE2E1</i>

Figures

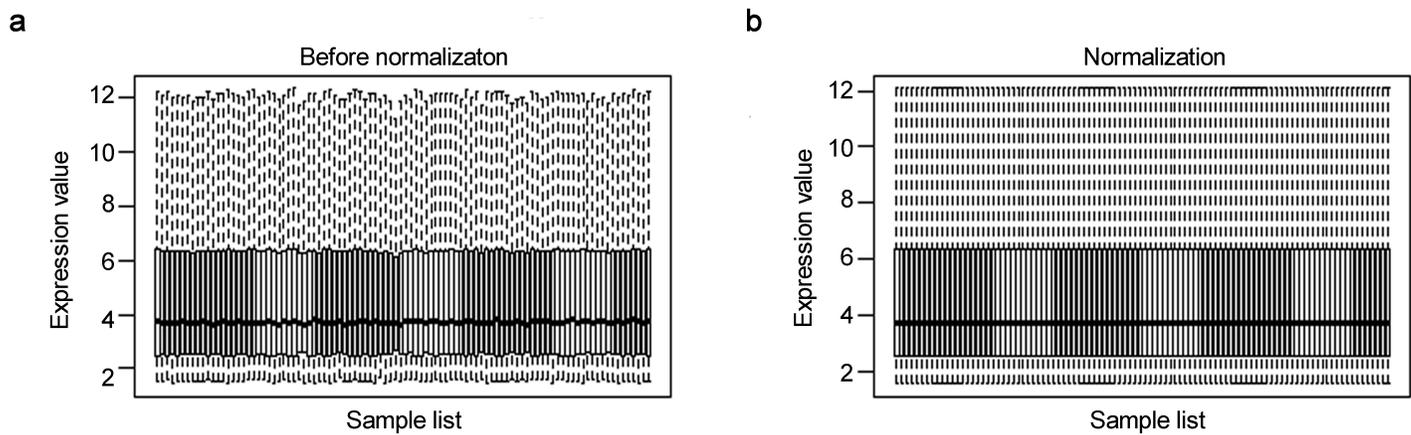


Figure 1

Normalization of the GSE44076 data set. a: Before normalization of expression value. b: Normalization of expression value.

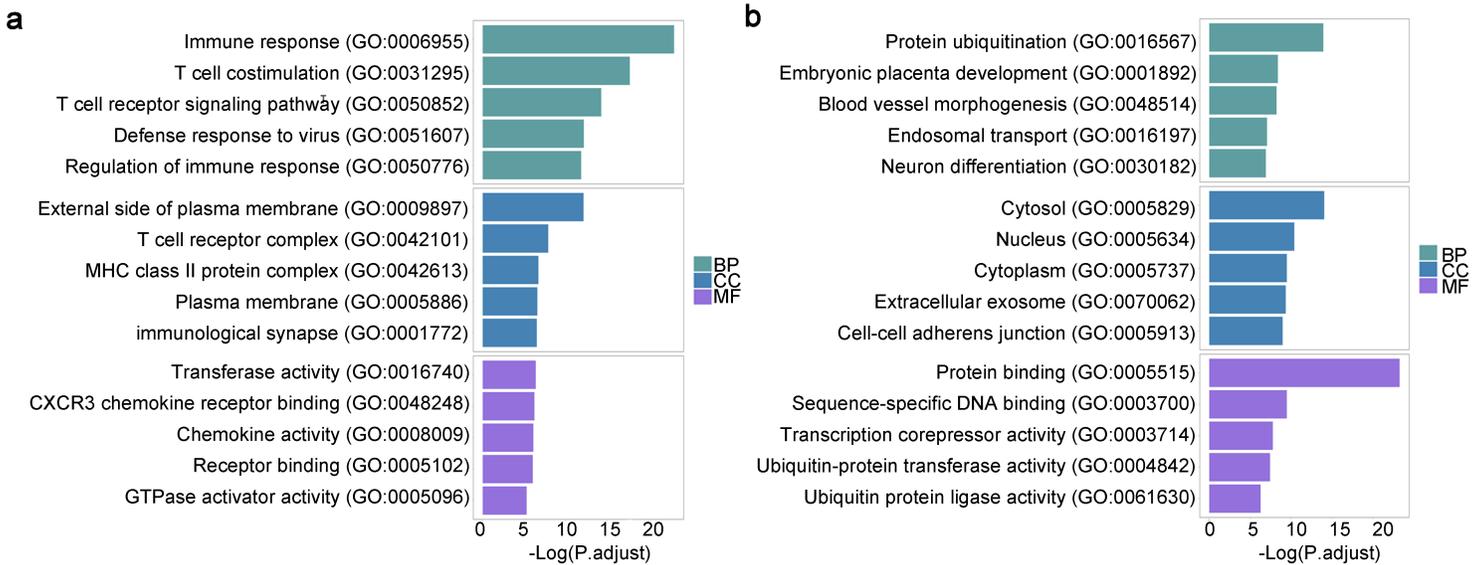


Figure 2

Significantly enriched term of GO analysis in colon cancer. a: Upregulated genes in RCC with the top 15 enriched GO terms. b: Upregulated genes in LCC with the top 15 enriched GO terms.

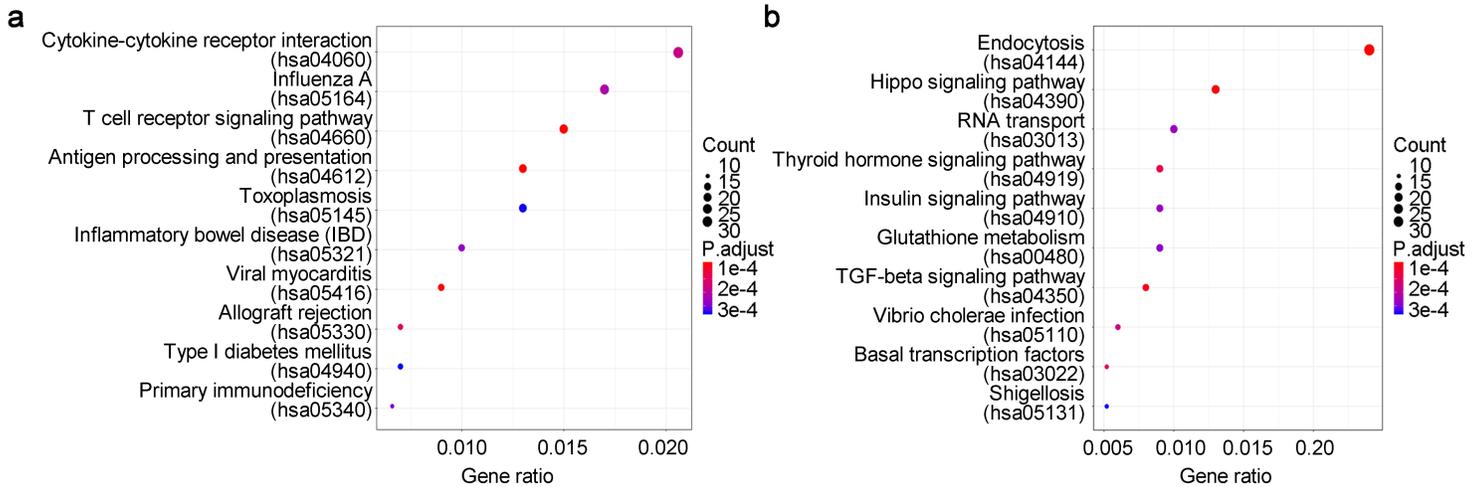


Figure 3

KEGG analysis. a: Upregulated genes in RCC with the top 10 enriched pathways. b: Upregulated DEGs in LCC with the top 10 enriched pathways.

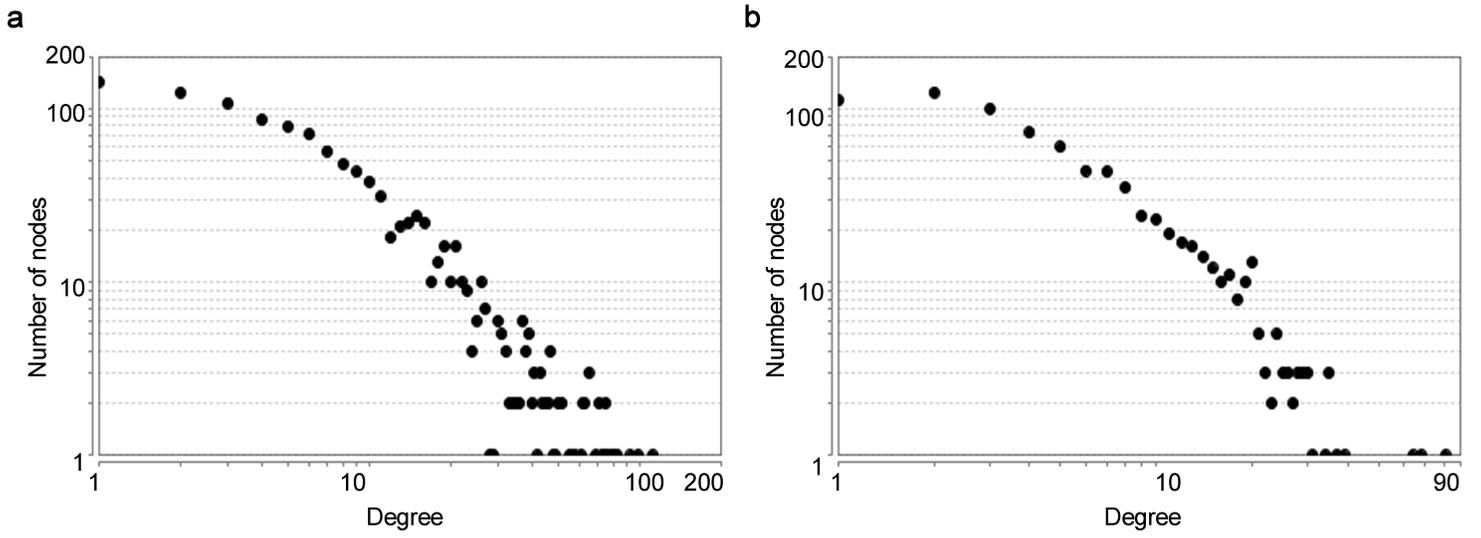


Figure 4

Nodes degree distributions in PPI networks. a: Nodes distribution in RCC network. b: Nodes distribution in LCC network.

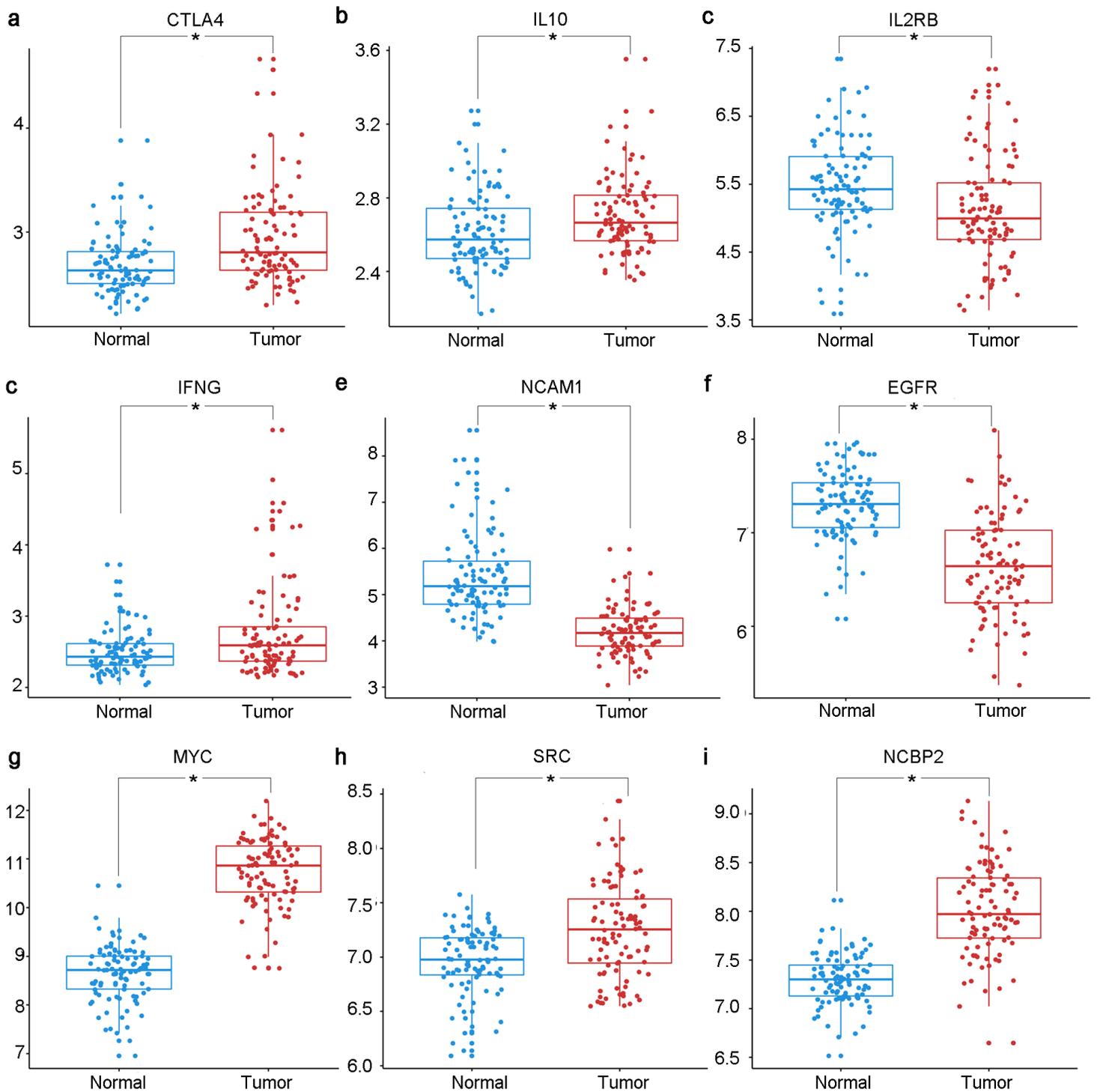


Figure 5

The genes expression levels of hub genes in GSE44076. The gene expression levels of CTAL4, IL10, IFNG, MYC, SRC, and NCBP2 were up-regulated in colon cancer samples compared to normal tissues, whereas IL2RB, NCAM1, and EGFR were decreased in cancer patients. a: CTAL4. b: IL10. c: IL2RB. d: IFNG. e: NCAM1. f: EGFR. g: MYC. h: SRC. i: NCBP2. *P < 0.01.

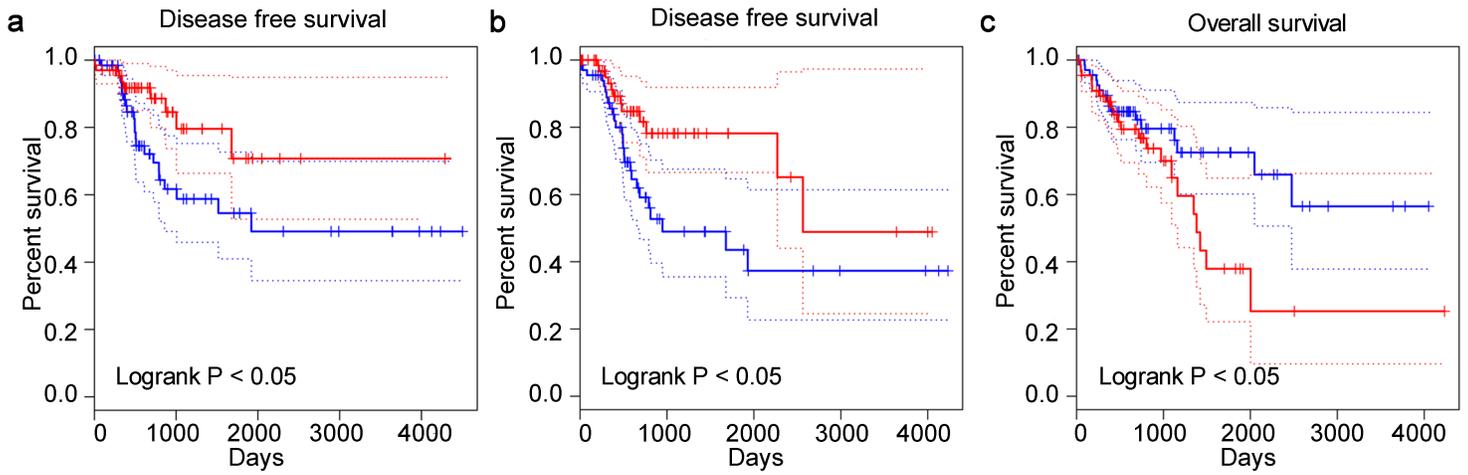


Figure 6

Survival analysis for hub genes in colon cancer patients. a: EGFR. b: IFNG. c: NCBP2. Red indicates relative gene high-expression cohorts; blue indicates the relative gene low-expression cohorts. The 95% confidence intervals were presented as dotted lines on the plot.