Bioinformatics Analysis of Key Genes and Pathways in Colorectal Cancer

YUEWEN QI,[1] HAOWEN QI,[2] ZEYUAN LIU,[3] PEIYUAN HE,[1,*] and BINGQING LI[1,*]

ABSTRACT

Colorectal cancer (CRC) is the third most prevalent cancer in the world. Although great progress has been made, the specific molecular mechanism remains unclear. This study aimed to explore the differentially expressed genes (DEGs) and underlying mechanisms of CRC using bioinformatics analysis. In this study, we identified a total of 1353 DEGs in the database of GSE113513, including 715 up- and 638 downregulated genes. Gene ontology analysis results showed that upregulated DEGs were significantly enriched in cell division, cell proliferation, and DNA replication. The downregulated DEGs were enriched in immune response, relation of cell growth and inflammatory response. The Kyoto Encyclopedia of Genes and Genomes pathway analysis showed that upregulated DEGs were enriched in cell cycle and p53 signaling pathway, whereas the downregulated DEGs were enriched in drug metabolism, metabolism of xenobiotics by cytochrome P450, and nitrogen metabolism. A total of 124 up-key genes and 35 down-key genes were identified from the protein–protein interaction networks. Furthermore, we identified five up-modules (up-A, up-B, up-C, up-D, and up-E) and three down-modules (d-A, d-B, and d-C) by module analysis. The module up-A was enriched in sister chromatid cohesion, cell division, and mitotic nuclear division. Pathways associated with cell cycle, progesterone-mediated oocyte maturation, oocyte meiosis, and p53 signaling pathway. Whereas the d-A was mainly enriched in G-protein coupled receptor signaling pathway, cell chemotaxis, and chemokine-mediated signaling pathway. The pathways enriched in chemokine signaling pathway, cytokine–cytokine receptor interaction, and alcoholism. These key genes and pathways might be used as molecular targets and diagnostic biomarkers for the treatment of CRC.

Keywords: colorectal cancer, key genes, molecular mechanisms.

1. INTRODUCTION

C OLORECTAL CANCER (CRC) is the third most common cancer and the second leading cause of cancer-related death in the world. Over 1.8 million new CRC cases and 881,000 deaths are estimated to occur in 2018, accounting for about 1 in 10 cancer cases and deaths (Bray et al., 2018). Although colonoscopy has

---

[1]Department of Gastroenterology, Affiliated Hospital of Chengde Medical College, Chengde, P.R. China.
[2]Department of Acupuncture and Massage, Chengde Hospital of Traditional Chinese Medicine, Chengde, P.R. China.
[3]Department of Special Medicine, Qingdao University Medical College, Qingdao, P.R. China.
*These authors' contributed equally to this work.

made great achievements in the diagnosis and treatment of early CRC, there is still a low rate of early CRC diagnosis, advanced CRC is extremely difficult to treat, and the prognosis of the patients is poor. The survival rate of advanced CRC is still <10% (Ciani et al., 2015; Moriarity et al., 2016). Therefore, it is urgent to investigate potential mechanisms and the targets to provide new ideas and indicators for clinical diagnosis and treatment of CRC.

The risk factors of CRC include smoking, obesity, alcohol, older age, high-fat intake, and lack of physical exercise (Zauber et al., 2012). Increasing evidence has indicated that the occurrence and development of CRC are a complex biological process (BP), which involves the abnormal expression of multiple genes and the functional changes of multiple pathways. It is known that the level of CLC was correlated with eosinophil density in inflammation and IFNAR1 was associated with oncogenesis. Ågesen et al. identified that CLC and IFNAR1 are associated with early onset CRC and provided additional support to the importance of the discriminated genes in the development of CRC within a genetic susceptibility context (Agesen et al., 2011). In the exon level study, Aziz et al. (2014) reported that ATP8B1 was a novel target for CRC. Xu et al. (2014) found that microRNA-375 was matched in CRC and associated with some critical signal pathways, such as Wnt, MAPK, and TGF-β. Although great progress have been made, the molecular mechanisms of CRC are far from been fully understood.

Analysis of abnormal genes and their biological pathways at genomic level can provide a more comprehensive understanding of key functional genes and their core mechanisms in CRC. Recently, gene expression profile data associated with CRC have been studied by many researchers. Jun Peng et al. constructed the microarray data of GSE113513, and identified the differentially expressed genes (DEGs) between 14 CRC tissues and 14 matched adjacent normal tissues by using the affymetrix human GeneChip primeview array. However, they did not perform further study of DEGs, such as the construction of a protein–protein interaction (PPI) network and module analysis of the PPI network.

In this study, we download the database of GEO113513 and identified the DEGs by comparing RNA expression from 14 CRC tissues and 14 matched adjacent normal tissues. Based on the DEGs, the gene ontology (GO) analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enrichment analyses were applied, and the PPI network and module were also constructed. In addition, we reanalyze the GO analysis and KEGG enrichment pathways of the module of PPI network. The findings of this study may provide new insights into CRC diagnosis and treatment.

2. MATERIALS AND METHODS

2.1. Data sets

The gene expression omnibus (GEO) is a public functional genomics data repository, which is freely available for users (Barrett et al., 2013). The GSE113513 gene expression profiles were downloaded from the GEO database. This data set was deposited by Jun Peng et al., which is based on the Affymetrix GPL15207 platform and has 28 samples, including 14 CRC tissues and 14 matched adjacent normal tissues.

2.2. Data processing

We used the packages of R statistical software to analyze the gene expression profile. The original CEL file data of the GSE113513 were processed into expression estimates, and background adjustment, normalization, and log transformation of the values were analyzed by affy package (Gautier et al., 2004). The probe-level data was transformed and matched with the gene symbols; when several probes correspond to one gene symbol, the gene expression value was the mean of the probes. The missing values of the genes were supplemented using the K-Nearest Neighbor. Furthermore, the LIMMA package was used to identify DEGs between CRC and matched adjacent normal tissues. The $\log_2$-fold change ($\log_2$FC) was calculated. We selected DEGs with $|\log_2$FC$| > 1$ and $p < 0.05$ in microarray data for further analysis. Then, we constructed heatmap and volcano for DEGs using the packages in R software.

2.3. GO and pathway enrichment analysis of DEGs

GO analysis is a common genes analysis method, which can provide functional classification for genomic data, including categories of BPs, cellular component (CC), and molecular function (MF) (Ashburner et al., 2000). The KEGG database is a knowledge bioinformatics resource for systematic analysis, annotation, and visualization of gene functions (Du et al., 2014). The Database for Annotation,

Visualization and Integrated Discovery (DAVID) is an online tool for gene functional classification, which can systematic and integrative analysis of large gene lists (Huang da et al., 2009). In this study, to analyze the functions of DEGs, GO enrichment and KEGG pathway analysis were conducted using the DAVID online tool; $p < 0.01$ was set as the cutoff point.

2.4. Integration of PPI network and key genes identification analysis

The Search Tool for the Retrieval of Interacting Genes (STRING) is a biological database designed to predict PPI information (von Mering et al., 2003). The DEGs were mapped to STRING to evaluate the interactive relationships, with a confidence score >0.9 defined as significant. Then, we use Cytoscape, a biological graph visualization tool for integrated models of biologic molecular interaction networks software, to construct PPI networks (Smoot et al., 2011). The key genes were identified with a connectivity degree >10.

2.5. Module analysis of the PPI network

The Molecular Complex Detection (MCODE), a plugin for Cytoscape, was used to screen the modules of the PPI network. The criteria were set as follows: degree cutoff = 2, node score cutoff = 0.2, k-core = 2, and maximum depth = 100 (Rhrissorrakrai and Gunsalus, 2011). Moreover, the biological function and pathway enrichment analysis were performed for key genes in the modules.

3. RESULT

3.1. Identification of DEGs

The raw data and the annotation information of the GSE113513 were downloaded from GEO database. Based on the LIMMA package analysis, a total of 1353 DEGs between 14 CRC samples and 14 matched adjacent normal samples were identified, of which 715 were upregulated and 638 were downregulated.

TABLE 1. GENE ONTOLOGY ANALYSIS OF DIFFERENTIALLY EXPRESSED GENES
ASSOCIATED WITH COLORECTAL CANCER

Category	Term	Description	Count	p
Upregulated				
CC	GO:0005730	Nucleolus	91	3.94E-22
CC	GO:0005654	Nucleoplasm	177	9.66E-18
CC	GO:0005634	Nucleus	257	2.94E-10
BP	GO:0051301	Cell division	47	1.63E-14
BP	GO:0007067	Mitotic nuclear division	39	2.51E-14
BP	GO:0006260	DNA replication	29	1.12E-12
BP	GO:0008283	Cell proliferation	39	3.51E-09
BP	GO:0006364	rRNA processing	28	1.17E-08
MF	GO:0005515	Protein binding	383	1.32E-09
MF	GO:0044822	poly(A) RNA binding	78	1.82E-08
Downregulated				
CC	GO:0005615	Extracellular space	106	4.67E-19
CC	GO:0070062	Extracellular exosome	156	9.97E-14
CC	GO:0005576	Extracellular region	104	8.23E-13
CC	GO:0005578	Proteinaceous extracellular matrix	25	3.57E-06
CC	GO:0031012	Extracellular matrix	26	6.37E-06
BP	GO:0006954	Inflammatory response	32	1.11E-06
BP	GO:0001558	Regulation of cell growth	13	6.70E-06
BP	GO:0006955	Immune response	32	9.67E-06
MF	GO:0008201	Heparin binding	26	1.78E-11
MF	GO:0005102	Receptor binding	30	1.41E-06

Category: GO function; count, the no. of DEGs; The *p*-value was adjusted with the Benjamini–Hochberg method.

BP, biological process; CC, cellular component; DEG, differentially expressed gene; GO, gene ontology; MF, molecular function.

TABLE 2. THE TOP FIVE ENRICHED KYOTO ENCYCLOPEDIA OF GENES AND GENOMES PATHWAYS
IN UPREGULATED DIFFERENTIALLY EXPRESSED GENES AND DOWNREGULATED
DIFFERENTIALLY EXPRESSED GENES

Category	Term	Description	Count	p
Upregulated				
KEGG	hsa04110	Cell cycle	24	8.49E-11
KEGG	hsa03008	Ribosome biogenesis in eukaryotes	18	1.19E-08
KEGG	hsa04115	p53 signaling pathway	11	1.48E-04
KEGG	hsa04914	Progesterone-mediated oocyte maturation	12	3.15E-04
KEGG	hsa03440	Homologous recombination	7	5.43E-04
Downregulated				
KEGG	hsa04978	Mineral absorption	10	2.02E-05
KEGG	hsa04610	Complement and coagulation cascades	10	7.39E-04
KEGG	hsa04060	Cytokine–cytokine receptor interaction	20	0.001110796
KEGG	hsa00910	Nitrogen metabolism	5	0.002668117
KEGG	hsa00982	Drug metabolism—cytochrome P450	9	0.00282832

Category: KEGG pathway; Term: *Homo sapiens*; Count, the no. of DEGs.
KEGG, Kyoto Encyclopedia of Genes and Genomes.

3.2. GO function enrichment analysis

To further investigate the functions of identified DEGs, all DEGs were uploaded in the DAVID software to analyze the GO categories and KEGG pathways. GO function result showed that the upregulated DEGs were markedly enriched in 22 terms and the downregulated DEGs were enriched in 14 terms. Among them, the upregulated DEGs were mainly enriched in nucleolus and nucleoplasm in CC, and the downregulated DEGs were markedly enriched in extracellular exosome, extracellular matrix, and proteinaceous extracellular matrix. GO BP analysis showed that upregulated DEGs were enriched in cell division, DNA replication, and cell proliferation, the downregulated DEGs were enriched in inflammatory response, immune response, and regulation of cell growth. For MF, the upregulated DEGs were mostly enriched in protein and poly(A) RNA binding, and downregulated DEGs were enriched in heparin and receptor binding. The top 10 GO terms in each are shown in Table 1 (all $p < 0.02$).

3.3. KEGG pathway analysis

We performed pathway enrichment analysis using the KEGG analysis of DAVID. The results showed that the upregulated DEGs were mainly enriched in nine KEGG pathways, including cell cycle, ribosome

TABLE 3. THE TOP 20 DIFFERENTIALLY EXPRESSED KEY GENES IN THE UPREGULATED
AND DOWNREGULATED PROTEIN–PROTEIN INTERACTION NETWORK

ID	Degree	ID	Degree	ID	Degree	ID	Degree
Upregulated							
CDK1	68	BUB1	42	BOP1	33	KIF11	31
CCNB1	57	MAD2L1	41	TOP2A	32	CHEK1	31
CCNB2	49	NOP58	38	FBL	32	RRP9	31
NOP56	43	KIF2C	33	POLR1B	32	UTP15	31
CCNA2	42	CENPE	33	NUP37	31	CIRH1A	31
Downregulated							
KNG1	35	GNAI1	22	NPY	20	CCL28	19
GNG8	28	CXCL12	21	NPY1R	20	CCL5	19
GNG7	28	SST	21	P2RY14	19	CCL21	19
LPAR1	24	NPBWR2	20	GPR18	19	CXCL13	19
C3	24	PYY	20	INSL5	19	CCL19	19

Degree: the connectivity of the protein.
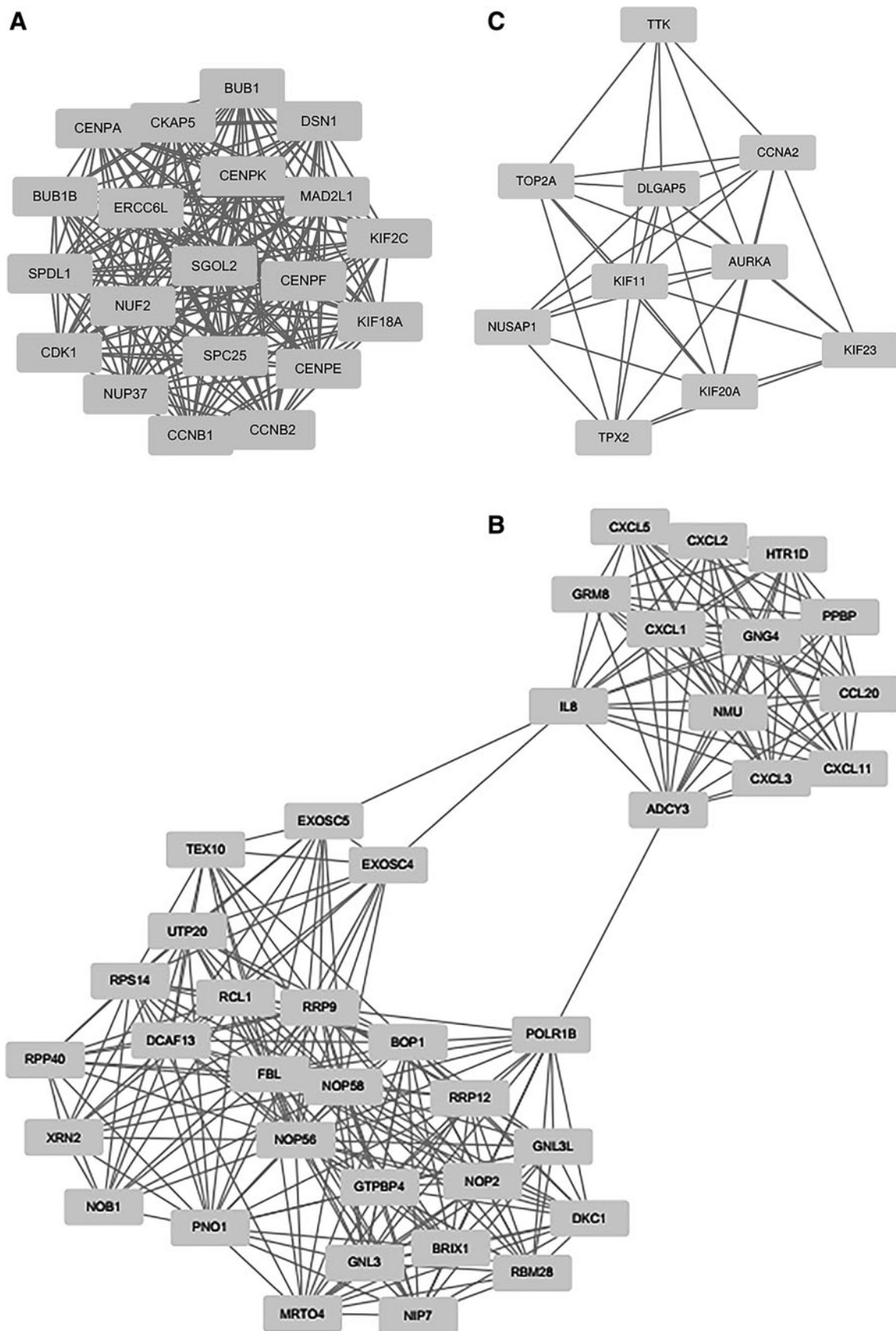
FIG. 1. Upregulated PPI network with MCODE score ≥4 and node ≥6. PPI, protein–protein interaction.
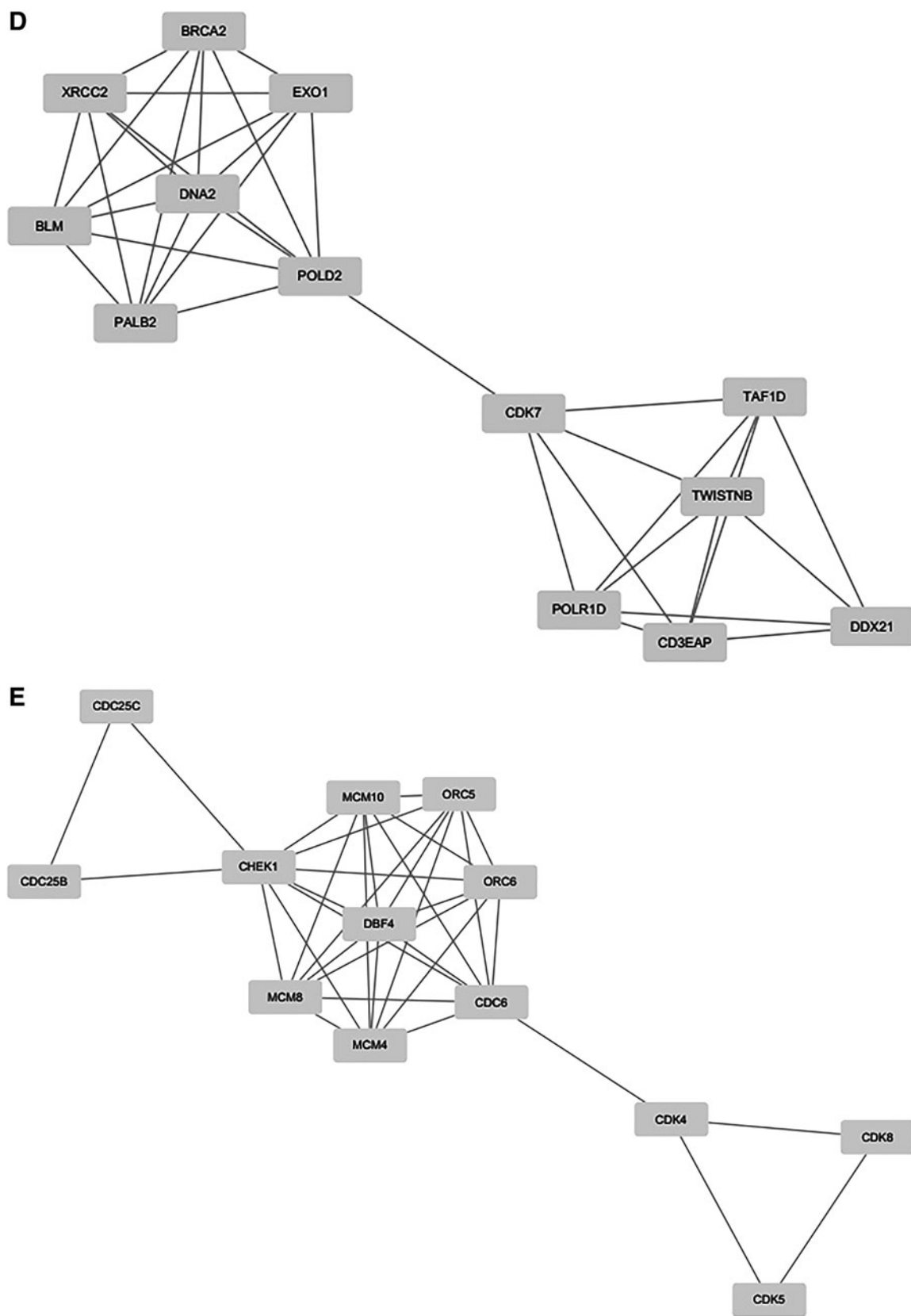
FIG. 1. Continued.

biogenesis in eukaryotes, p53 signaling pathway, and progesterone-mediated oocyte maturation, whereas the downregulated DEGs were enriched in mineral absorption, complement, and coagulation cascades, cytokine–cytokine receptor interaction, nitrogen metabolism, and the drug metabolism of cytochrome P450 (all $p < 0.01$). The top five enriched KEGG pathways in each are shown in Table 2.

3.4. PPI network construction and hub gene identification

After the upregulated and downregulated DEGs were submitted into STRING, a total of 1657 upregulated and 470 downregulated PPI relationships were obtained. Then, the key genes in the networks with connectivity degree >10 were identified. A total of 124 key genes were selected from the upregulated PPI network, which included *CDK1*, *CCNB1*, and *CCNB2*. Meanwhile, 35 key genes, such as *KNG1*, *GNG8*, and *GNG7*, were identified from the downregulated PPI network. The top 20 key genes in each are shown in Table 3.

3.5. Module analysis

Five modules with MCODE score ≥4 and nodes ≥6 were selected in the upregulated PPI networks: up-A (MCODE score = 20) with 20 nodes, up-B (MCODE score = 15,128) with 40 nodes, up-C (MCODE score = 8.222) with 10 nodes, up-D (MCODE score = 6) with 13 nodes, and up-E (MCODE score = 5.833) with 13 nodes (Fig. 1). Meanwhile, there are three modules in the downregulated PPI network: d-A
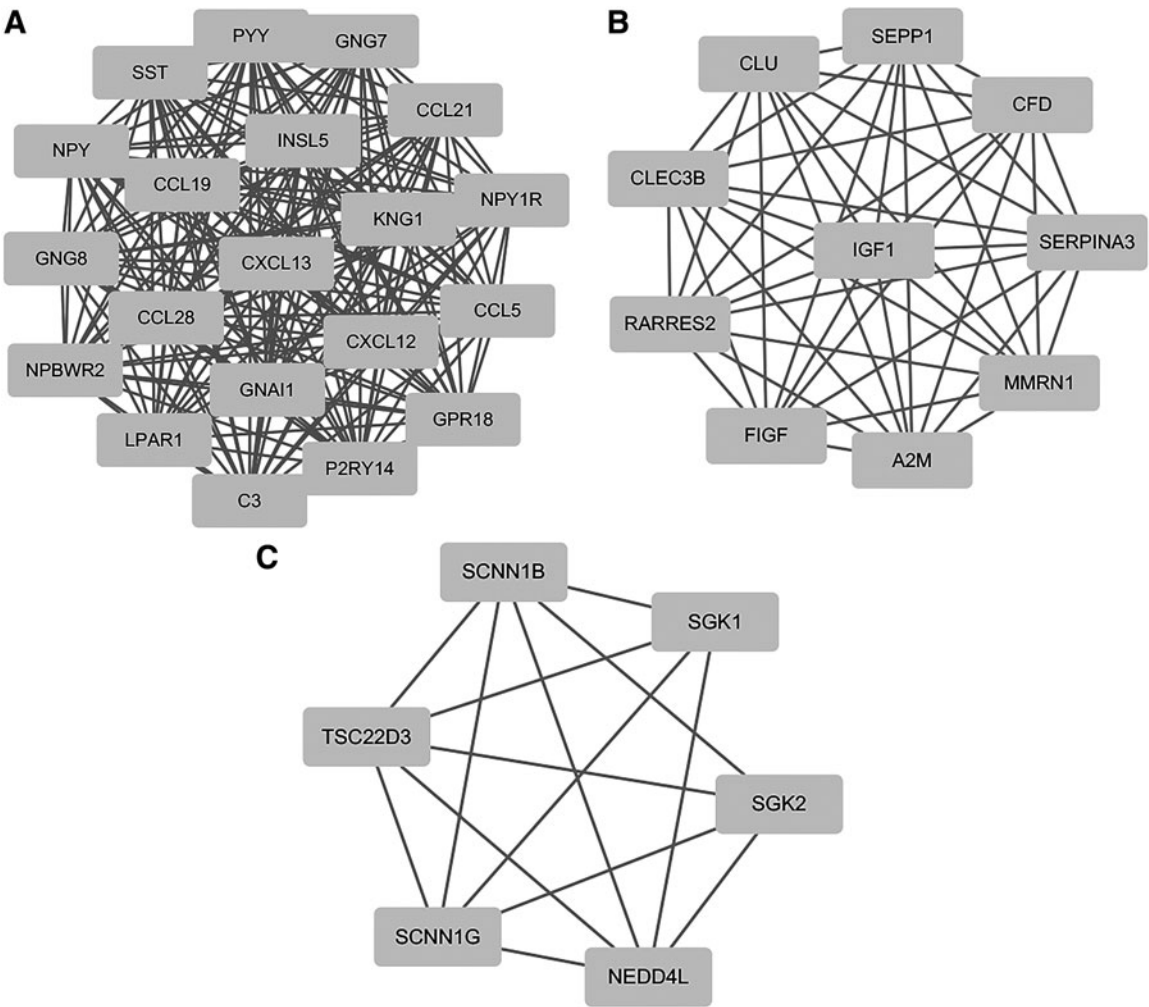

FIG. 2. Downregulated PPI network with MCODE score ≥4 and node ≥6.

TABLE 4. THE GENE ONTOLOGY FUNCTIONAL ENRICHMENT ANALYSIS OF DIFFERENTIALLY
EXPRESSED GENES IN THE UP-A AND DOWN-A MODULES

Category	Term	Description	Count	p
Upregulated				
BP	GO:0007062	Sister chromatid cohesion	16	1.81E-31
BP	GO:0051301	Cell division	16	3.47E-23
BP	GO:0007067	Mitotic nuclear division	11	1.63E-14
BP	GO:0007080	Mitotic metaphase plate congression	5	5.97E-08
CC	GO:0000777	Condensed chromosome kinetochore	11	1.52E-19
CC	GO:0000776	Kinetochore	9	4.49E-15
CC	GO:0005829	Cytosol	19	4.58E-14
CC	GO:0000940	Condensed chromosome outer kinetochore	4	4.52E-08
CC	GO:0000922	Spindle pole	6	5.62E-08
CC	GO:0000775	Chromosome, centromeric region	5	2.55E-07
Downregulated				
BP	GO:0007186	G-protein coupled receptor signaling pathway	15	1.30E-14
BP	GO:0060326	Cell chemotaxis	6	8.28E-09
BP	GO:0070098	Chemokine-mediated signaling pathway	5	1.08E-06
BP	GO:0006954	Inflammatory response	7	2.69E-06
BP	GO:0006955	Immune response	7	4.93E-06
BP	GO:1903237	Negative regulation of leukocyte tethering or rolling	3	7.27E-06
CC	GO:0005576	Extracellular region	12	9.59E-08
CC	GO:0005615	Extracellular space	11	2.34E-07
MF	GO:0008009	Chemokine activity	6	1.88E-09
MF	GO:0042379	Chemokine receptor binding	4	6.75E-08

Category: GO function; ; Count, the no. of DEGs. The *p*-value was adjusted with the Benjamini–Hochberg method.

(MCODE score = 20) with 20 nodes, d-B (MCODE score = 10) with 10 nodes, and d-C (MCODE score = 5.6) with 6 nodes. The top one module in each is shown in Figure 2.

The GO functional enrichment scores of modules up-A were mainly enriched in sister chromatid cohesion, cell division, mitotic nuclear division, and mitotic nuclear division; module d-A was enriched in G-protein coupled receptor signaling pathway, cell chemotaxis, chemokine-mediated signaling pathway, inflammatory response, and immune response (Table 4). In addition, the KEGG pathway analysis revealed that the DEGs in module up-A were markedly enriched in cell cycle, progesterone-mediated oocyte maturation, oocyte meiosis, and p53 signaling pathway; the DEGs in module d-A were significantly enriched in pathways including chemokine signaling pathway, cytokine–cytokine receptor interaction, alcoholism, and the regulation of lipolysis in adipocytes (Table 5).

TABLE 5. THE KYOTO ENCYCLOPEDIA OF GENES AND GENOMES PATHWAYS ANALYSIS
OF DIFFERENTIALLY EXPRESSED GENES IN THE UPREGULATED-A AND DOWNREGULATED-A
MODULES WITH THE THRESHOLD OF *P* < 0.01

Category	Term	Description	Count	p
Upregulated				
KEGG	hsa04110	Cell cycle	6	1.04E-08
KEGG	hsa04914	Progesterone-mediated oocyte maturation	5	3.51E-07
KEGG	hsa04114	Oocyte meiosis	5	9.40E-07
KEGG	hsa04115	p53 signaling pathway	3	0.001366951
Downregulated				
KEGG	hsa04062	Chemokine signaling pathway	9	2.63E-09
KEGG	hsa04060	Cytokine–cytokine receptor interaction	6	1.67E-04
KEGG	hsa04923	Regulation of lipolysis in adipocytes	3	0.007260181
KEGG	hsa05034	Alcoholism	4	0.007327022

Category: KEGG pathway; Term: *Homo sapiens*; Count: the no. of DEGs.

4. DISCUSSION

In this study, we applied a bioinformatics approach to investigate the DEGs in patients with CRC tissues and matched adjacent normal tissues. Based on DEGs data, the molecular mechanism of CRC was identified by DEGs gene ontology and pathway enrichment analysis and PPI network construction. Then we identified the key functional genes and investigated the core module using functional and enrichment analysis. These results are strong supplement for further understanding of CRC diagnosis and treatment.

Our DEGs screening results identified 1353 DEGs between 14 CRC samples and 14 matched adjacent normal samples, which included 715 upregulated and 638 downregulated DEGs. These upregulated DEGs were obviously enriched in cell proliferation and protein and poly(A) RNA binding, whereas the downregulated DEGs were enriched in inflammatory response, immune response, and regulation of cell growth. All these are compatible with the important mechanism of cancer occurrence and development (Diakos et al., 2014; Kolch et al., 2015; Chiou and Burotto, 2015; Chung et al., 2016; Fridman et al., 2017; Hojman, 2017).

KEGG pathway enrichment analysis of upregulated DEGs identified a total of 21 pathways, including cell cycle, ribosome biogenesis in eukaryotes, p53 signaling pathway, and homologous recombination. Disruptions of the cell cycle have been well documented to be involved in the genesis and propagation of a variety of cancers. In addition, several aberrations in the cell cycle checkpoints have been shown to have prognostic significance in CRC (McKay et al., 2002; Santamaria and Ortega, 2006; Solier et al., 2012; Balakrishnan et al., 2016). Iacopetta et al. revealed that p53 mutations that lose transactivational ability are more common in advanced CRC and associated with poor survival (Iacopetta et al., 2006). Meanwhile, the impaired ribosome biogenesis is also to be considered a risk factor for cancer initiation (Goudarzi and Lindstrom, 2016).

In addition, the downregulated DEGs pathway enrichment analysis identified 25 pathways, such as complement and coagulation cascades, nitrogen metabolism, and cytochrome P450 of drug metabolism. In agreement with our study, after upregulation of complement C3a receptor on neutrophils and activation of the complement cascade by increasing circulating lipopolysaccharide, Guglietta et al. (2016) revealed that tumorigenesis is associated with coagulation, neutrophilia, and complement activation. In liver cancer study, the aberrant expression of urea cycle enzymes was found to maximize the use of nitrogen and carbon for the anabolic synthesis of macromolecules that are required during tumor proliferation and growth (Keshet et al., 2018). Several cytochrome P450s may act as independent markers of CRC prognosis (Kumarakulasingham et al., 2005).

Our PPI network analysis constructed with DEGs identified key genes including *CDK1*, *FBL*, *CIRH1A*, *KIF2C*, *TOP2A*, *KNG1*, *LPAR1*, *GNG7*, *CCL5*, and *GNAI1*. In addition, a modular analysis of the PPI network indicates that the CRC is primarily associated with eight core modules, including five upregulated cores and three downregulated cores. Based on these results, we further explored the biological function and pathway enrichment for up-A and d-A modules. The up-A results showed that CRC was associated with sister chromatid condensation, cell division, mitotic nuclear division, and mitotic midplate sessions. The d-A module shows that CRC is involved in G protein-coupled receptor signaling pathways, cell chemotaxis, chemokine-mediated signaling pathways, inflammatory responses, immune responses, and negative regulation of leukocyte binding or rolling.

Through literature mining of those key genes in up-A module, we have found that a large number of studies on the CDK family including CDK1 are closely related to the development of CRC (He et al., 2018.). Small nucleolar RNAs are emerging as a novel class of proto-oncogenes and tumor suppressors. Nucleolar protein 5A (NOP56), nucleolar protein 5 (NOP58), and fibrillarin (FBL) as the three core proteins of box C/D small nucleolar ribonucleoprotein complexes play an essential role in ribosome assembly. Qin et al. (2017) reported that FBL regulated rRNA methylation patterns and promoted tumorigenesis. However, their involvement in tumorigenesis remains unclear. The other two key genes of *CIRH1A* and *KIF2C* were reported that significantly overexpressed in CRC. Gnjatic et al. reported that *KIF2C* induces frequent T cell response in patients with CRC (Gnjatic et al., 2010; Guo et al., 2017). Another key gene, *TOP2A*, functions as the target for several anticancer agents, and a variety of mutations in this gene have been associated with the development of drug resistance. In CRC, increased topoisomerase IIalpha expression is associated with advanced disease and chemotherapeutic resistance through inhibition of apoptosis (Coss et al., 2009).

The other five hub genes are *KNG1*, *LPAR1*, *GNG7*, *CCL5*, and *GNAI1*. It has been found that *KNG1* was shown to be a potential serum predictor of advanced colorectal adenoma and cancer (Wang et al., 2013).

Florence Quesada-Calvo et al. also reported that OLFM4 and KNG1 are abnormally expressed in CRC tissues, especially in early pTNM CRC stages. This could also be observed but more moderately in precancerous lesions. Further studies on these proteins are warranted to better understand these roles in CRC progression and their potential as early diagnostic tools (Quesada-Calvo et al., 2017). *LPAR1* was reported to regulate cell proliferation, survival, angiogenesis, or migration (Wei et al., 2013). Further studies are needed to determine the role of *LPAR1* in CRC. For *GNAI1*, G protein subunit alpha i1, which participates in many physiological processes, including adhesion, proliferation, and differentiation, however, the biological functions of *GNAI1* in CRC remain unclear (Foley et al., 2010). Another key gene, *GNG7*, belongs to a family of large G-proteins that may be involved in cell-contact-induced growth arrest and function in tumor suppression. *GNG7* was downregulated in gastrointestinal cancer, and *GNG7* suppression represents a new prognostic indicator in cases of esophageal cancer (Ohta et al., 2008), To date, the biological functions of *GNG7* in CRC remain unclear.

After biological functions and pathway enrichment analysis, the up-A results showed that upregulated key genes were associated with cell cycle and cell division, which play an essential role in proliferation and growth in CRC. The d-A module shows that downregulated key genes were involved in G protein-coupled receptor signaling pathways, cell chemotaxis, chemokine-mediated signaling pathways, inflammatory responses, immune responses, and negative regulation of leukocyte binding or rolling. These pathways were associated with the processes in the prognosis and treatment of CRC. To investigate the detailed mechanism, further studies are needed to be verified.

In conclusion, our study is based on the study of Jun Peng et al., which lays the foundation for further research on the underlying molecular mechanisms of CRC. These findings identified the key genes and pathways of CRC. However, this study also has some limitations, such as the lack of experiments and small number of samples. Therefore, further experimental studies based on our findings are needed.

ACKNOWLEDGMENT

AUTHOR DISCLOSURE STATEMENT

The authors declare there are no competing financial interests.

REFERENCES

Agesen, T.H., Berg, M., Clancy, T., et al. 2011. CLC and IFNAR1 are differentially expressed and a global immunity score is distinct between early- and late-onset colorectal cancer. *Genes Immun.* 12, 653–662.

Ashburner, M., Ball, C.A., Blake, J.A., et al. 2000. Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nat. Genet.* 25, 25–29.

Aziz, M.A., Periyasamy, S., Al Yousef, Z., et al. 2014. Integrated exon level expression analysis of driver genes explain their role in colorectal cancer. *PLoS One* 9, e110134.

Balakrishnan, A., Vyas, A., Deshpande, K., et al. 2016. Pharmacological cyclin dependent kinase inhibitors: Implications for colorectal cancer. *World J. Gastroenterol.* 22, 2159–2164.

Barrett, T., Wilhite, S.E., Ledoux, P., et al. 2013. NCBI GEO: Archive for functional genomics data sets—Update. *Nucl. Acids Res.* 41, D991–D995.

Bray, F., Ferlay, J., Soerjomataram, I., et al. 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 68, 394–424.

Chiou, V.L., and Burotto, M. 2015. Pseudoprogression and immune-related response in solid tumors. *J. Clin. Oncol.* 33, 3541–3543.

Chung, W., Kim, M., de la Monte, S., et al. 2016. Activation of signal transduction pathways during hepatic oncogenesis. *Cancer Lett.* 370, 1–9.

Ciani, O., Buyse, M., Garside, R., et al. 2015. Meta-analyses of randomized controlled trials show suboptimal validity of surrogate outcomes for overall survival in advanced colorectal cancer. *J. Clin. Epidemiol.* 68, 833–842.

Coss, A., Tosetto, M., Fox, E.J., et al. 2009. Increased topoisomerase IIalpha expression in colorectal cancer is associated with advanced disease and chemotherapeutic resistance via inhibition of apoptosis. *Cancer Lett*. 276, 228–238.

Diakos, C.I., Charles, K.A., McMillan, D.C., et al. 2014. Cancer-related inflammation and treatment effectiveness. *Lancet Oncol*. 15, e493–e503.

Du, J., Yuan, Z., Ma, Z., et al. 2014. KEGG-PATH: Kyoto encyclopedia of genes and genomes-based pathway analysis using a path analysis model. *Mol. Biosyst*. 10, 2441–2447.

Foley, J.F., Singh, S.P., Cantu, M., et al. 2010. Differentiation of human T cells alters their repertoire of G protein alpha-subunits. *J. Biol. Chem*. 285, 35537–35550.

Fridman, W.H., Zitvogel, L., Sautes-Fridman, et al. 2017. The immune contexture in cancer prognosis and treatment. *Nat. Rev. Clin. Oncol*. 14, 717–734.

Gautier, L., Cope, L., Bolstad, B.M., et al. 2004. affy—Analysis of Affymetrix GeneChip data at the probe level. *Bioinformatics* 20, 307–315.

Gnjatic, S., Cao, Y., Reichelt, U., et al. 2010. NY-CO-58/KIF2C is overexpressed in a variety of solid tumors and induces frequent T cell responses in patients with colorectal cancer. *Int. J. Cancer* 127, 381–393.

Goudarzi, K.M., and Lindstrom, M.S. 2016. Role of ribosomal protein mutations in tumor development (Review). *Int. J. Oncol*. 48, 1313–1324.

Guglietta, S., Chiavelli, A., Zagato, E., et al. 2016. Coagulation induced by C3aR-dependent NETosis drives protumorigenic neutrophils during small intestinal tumorigenesis. *Nat. Commun*. 7, 11037.

Guo, F., Chen, J.J., and Tang, W.J. 2017. CIRH1A augments the proliferation of RKO colorectal cancer cells. *Oncol. Rep*. 37, 2375–2381.

He, L.J., Yang, D.L., Li, S.Q., et al. 2018. Facile construction of fused benzimidazole-isoquinolinones that induce cell-cycle arrest and apoptosis in colorectal cancer cells. *Bioorg. Med. Chem*. 26, 3899–3908.

Hojman, P. 2017. Exercise protects from cancer through regulation of immune function and inflammation. *Biochem. Soc. Trans*. 45, 905–911.

Huang da, W., Sherman, B.T., and Lempicki, R.A. 2009. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat.. Protoc*. 4, 44–57.

Iacopetta, B., Russo, A., Bazan, V., et al. 2006. Functional categories of TP53 mutation in colorectal cancer: Results of an International Collaborative Study. *Ann. Oncol*. 17, 842–847.

Keshet, R., Szlosarek, P., Carracedo, A., et al. 2018. Rewiring urea cycle metabolism in cancer to support anabolism. *Nat. Rev. Cancer* 18, 634–645.

Kolch, W., Halasz, M., Granovskaya, M., et al. 2015. The dynamic control of signal transduction networks in cancer cells. *Nat. Rev. Cancer* 15, 515–527.

Kumarakulasingham, M., Rooney, P.H., Dundas, S.R., et al. 2005. Cytochrome p450 profile of colorectal cancer: Identification of markers of prognosis. *Clin. Cancer Res*. 11, 3758–3765.

McKay, J.A., Douglas, J.J., Ross, V.G., et al. 2002. Analysis of key cell-cycle checkpoint proteins in colorectal tumours. *J. Pathol*. 196, 386–393.

Moriarity, A., O'Sullivan, J., Kennedy, J., et al. 2016. Current targeted therapies in the treatment of advanced colorectal cancer: A review. *Ther. Adv. Med. Oncol*. 8, 276–293.

Ohta, M., Mimori, K., Fukuyoshi, Y., et al. 2008. Clinical significance of the reduced expression of G protein gamma 7 (GNG7) in oesophageal cancer. *Br. J. Cancer* 98, 410–417.

Qin, W., Lv, P., Fan, X., et al. 2017. Quantitative time-resolved chemoproteomics reveals that stable O-GlcNAc regulates box C/D snoRNP biogenesis. *Proc. Natl Acad. Sci. U. S. A*. 114, E6749–e6758.

Quesada-Calvo, F., Massot, C., Bertrand, V., et al. 2017. OLFM4, KNG1 and Sec24C identified by proteomics and immunohistochemistry as potential markers of early colorectal cancer stages. *Clin. Proteomics* 14, 9.

Rhrissorrakrai, K., and Gunsalus, K.C. 2011. MINE: Module identification in networks. *BMC Bioinformatics* 12, 192.

Santamaria, D., and Ortega, S. 2006. Cyclins and CDKS in development and cancer: Lessons from genetically modified mice. *Front Biosci*. 11, 1164–1188.

Smoot, M.E., Ono, K., Ruscheinski, J., et al. 2011. Cytoscape 2.8: New features for data integration and network visualization. *Bioinformatics* 27, 431–432.

Solier, S., Zhang, Y.W., Ballestrero, A., et al. 2012. DNA damage response pathways and cell cycle checkpoints in colorectal cancer: Current concepts and future perspectives for targeted treatment. *Curr. Cancer Drug Targets* 12, 356–371.

von Mering, C., Huynen, M., Jaeggi, D., et al. 2003. STRING: A database of predicted functional associations between proteins. *Nucl. Acids Res*. 31, 258–261.

Wang, J., Wang, X., Lin, S., et al. 2013. Identification of kininogen-1 as a serum biomarker for the early detection of advanced colorectal adenoma and colorectal cancer. *PLoS One* 8, e70519.

Wei, J.S., Johansson, P., Chen, L., et al. 2013. Massively parallel sequencing reveals an accumulation of de novo mutations and an activating mutation of LPAR1 in a patient with metastatic neuroblastoma. *PLoS One* 8, e77731.

Xu, L., Li, M., Wang, M., et al. 2014. The expression of microRNA-375 in plasma and tissue is matched in human colorectal cancer. *BMC Cancer* 14, 714.

Zauber, A.G., Winawer, S.J., O'Brien, M.J., et al. 2012. Colonoscopic polypectomy and long-term prevention of colorectal-cancer deaths. *N. Engl. J. Med.* 366, 687–696.

Address correspondence to:
Prof. Bingqing Li
Department of Gastroenterology
Affiliated Hospital of Chengde Medical College
36 Nanyingzi Street
Chengde
Hebei
P.R. China

E-mail: libq68@sohu.com

Dr. Peiyuan He
Department of Gastroenterology
36 Nanyingzi Street
Affiliated Hospital of Chengde Medical College
Chengde
Hebei
P.R. China

E-mail: peiyuanhe@163.com