frontiers

Identifying functional evolution processes according to the pathological stages of colorectal cancer

First Author [1,*], **Co-Author** [2] **and Co-Author** [1,2]

[1] *Laboratory X, Institute X, Department X, Organization X, City X , State XX (only USA, Canada and Australia), Country X*

[2] *Laboratory X, Institute X, Department X, Organization X, City X , State XX (only USA, Canada and Australia), Country X*

Correspondence*:
Corresponding Author
email@uni.edu

ABSTRACT

Colorectal cancer (CRC) is one of the malignant tumors with high morbidity and mortality. A prevalent method for studying cancer is to identify differentially expressed genes (DEGs) between control and patient samples, and conduct the following functional enrichment analyses. However, many of those studies ignore the fact that different pathological stages of the cancer are often highly different from each other. The mixture of those heterogeneous samples may lack the efficiency of identifying the real DEGs, and loss the opportunity to analyze the dynamic evolution process of cancer. In this study, we develop a feasible frame to identify function evolution processes of cancers according to their pathological stages as follows. Firstly, the limma package was used to identify four DEG sets between control samples and CRC stage I, II, III, and IV samples, followed by modules cluster from the PPI network on four DEG sets. Secondly, four functional module networks were constructed by the relationship of modules within four stage respectively, and then a functional evolution network was generated by analyzing the relationship of modules in adjacent stages. Finally, functional enrichment analysis were performed for DEGs of close modules in the functional evolution network. There were 2 significant paths in the functional evolution network: one is a2-b1-c2-d1 (Mod1) enriched in 14 pathways, the other is a1-b2-c1-d2 (Mod2) enriched in 8 pathways. The significant pathways obtained in present paper may play critical roles in the development of CRC, which contributes to explore molecular mechanisms and evolution processes of CRC.

Keywords: colorectal cancer, functional module, functional evolution network, functional enrichment analysis, pathway

1 INTRODUCTION

Colorectal cancer (CRC), with high morbidity and high mortality, is one of the malignant tumors. The incidence and mortality of CRC are among the top four cancers. It is predicted that 145,600 patients will be diagnosed with CRC and 51,020 patients will die of this cancer in 2019 (Siegel et al., 2019). Therefore, exploring evolution processes of CRC is conducive to study the molecular mechanism of CRC, and also promote the development of prognostic markers and therapeutic methods.

In recent years, several cell signal transduction pathways related to CRC development have been discovered, including Wnt/-catenin (Klaus and Birchmeier, 2008; Shim et al., 2015; Oh et al., 2014), Hedgehog (You et al., 2010), TGF-/Smads (Xiong et al., 2002), PI3K/Akt (Rychahou et al., 2005), MAPK (Gulmann et al., 2009) and p53 (Li et al., 2015; Russo et al., 2005) signaling pathways. The signaling pathways are of great importance for studying the molecular mechanisms of CRC development. However, there are still many CRC-related signal pathways that have not been discovered.

A prevalent method for studying colorectal cancer is to identify differentially expressed genes (DEGs) between control and patient samples, followed by the protein-protein interactions (PPI) analyses. However, many of those studies ignore the fact that different pathological stages of the cancer are often highly different from each other. The mixture of those heterogeneous samples may lack the efficiency of identifying the real DEGs, and loss the opportunity to analyze the dynamic evolution process of cancer.

To solve the issue, we propose a new approach to study the functional evolution processes of the CRC development. Firstly, we identified four differentially expressed genes (DEGs) sets between control and four CRC stages samples, respectively, followed by the Protein-Protein Interactions (PPI) analysis. Then, modules in PPI networks were identified by cluster for function analysis. Finally, we constructed four functional module networks and a functional evolution network based on modules above, which was contributed to analyse functional evolution processes underlying the pathological stages of CRC.

2 MATERIALS AND METHOD

2.1 Data Source

The gene expression profile dataset was obtained from Gene Expression Omnibus (GSE62932, http://www.ncbi.nlm.nih.gov/geo/). It includes 64 CRC samples and 4 healthy controls. The number of samples of stage I, II, III, and IV in the four stages of CRC is 12, 17, 20, and 15, respectively.

2.2 Data Processing

Raw microarray data was preprocessed through the R language affy package(https://www.bioconductor.org/).The raw data were normalized with Robust Multichip Average(RMA) method (Irizarry et al., 2003; Gautier et al., 2004). After filtering out nonspecific probes, the rest probes were mapped to corresponding gene symbols using the annotation information package in GPL570, hgu133plus2.db (Carlson et al., 2016). When a gene corresponds to multiple probes, we use the average expression value of all probes as the final expression value of this gene. Finally, the gene expression profiles of 20,186 genes was obtained.

2.3 Differential Expression Analysis

Differentially expressed genes (DEGs) between CRC stage I, II, III, and IV samples and control samples were screened respectively by the limma (Ritchie et al., 2015; Smyth et al., 2005) algorithm. Using the Benjamini-Hochberg (Ferreira, 2007) method to correct p-value to acquire FDR. Calculating log2-Fold Change (log2FC) value of each gene between two groups. Only the genes that meets $FDR < 0.01$ and $|log2FC| > 1.5$ were defined as DEGs.

2.4 Protein-Protein Interactions (PPIs) Analyses

STRING (Search Tool for the Retrieval of Interacting Genes) (Szklarczyk et al., 2014) database, an online server for searching for known genes or protein interaction, was utilized to obtain the PPIs between DEGs identified above. In this paper, species was Homo, and PPI score criterion was 0.4. Four PPI networks were

constructed by PPIs of four DEGs, which were visualized by using Cytoscape software (Shannon et al., 2003).

2.5 Functional Module Analyses of PPI Networks

Proteins do not function independently, but interact with others to mediate signaling pathways, cellular processes, and biological systems. Therefor, PPI information can aggregate functionally related DEGs to a functional module, which is important to investigated the module's functions. Identifying modules' function in a PPI network is important for understanding the complex molecular mechanisms of cancer. ClusterONE (Nepusz et al., 2012) was employed to identify functional modules with overlap neighborhood expansion (criteria: Min size=3, overlap threshold=0.8, simalarity = coefficient) in four stages of PPI network, respectively. Some functional modules of a PPI network have common edges, so we can consider these functional modules may play a more critical roles. A function module network was constructed by these modules as vertices and the number of common PPIs between functional modules as the weight of the edge and visualized with Cytoscape (Shannon et al., 2003). Precisely, given two different functional modules p and q from same stage, we define P and Q as the protein sets of p and q, and the number of PPI between proteins in P and Q was regarded as the weight between p and q.

2.6 Functional Evolution Analyses

As the functional modules of PPI network at different stages may contain some overlapped proteins, these functional modules from various stages can be considered as the focus to analyse, which is of great significance to identify the functional evolution process of cancer at different stages. Hence, we can create a functional evolution network in terms of the relationship between functional modules from adjacent stages and visualized with Cytoscape (Shannon et al., 2003). In this network, all functional modules of the four stages as vertices, the number of identical genes of functional modules between adjacent stages as the weight of the edge. Precisely, for example, given two functional modules a and b from stage I and stage II, we define A and B as the genes sets of a and b, and the intersection of A and B was defined $S = A \cap B$ and the gene number of set S was set as the weight between a and b.

2.7 Functional Enrichment Analyses

DAVID (Database for Annotation, Visualization and Integrated Discovery) (Sherman et al., 2009) provides a comprehensive set of functional annotation tools for researchers to understand biological function information of large gene lists. GO (Gene Ontology) enrichment analysis and KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway enrichment analysis were performed for the DEGs of functional modules with DAVID. Over-represented GO categories or KEGG pathways were identified and $EASE < 0.05$ was set as the criterion.

3 RESULT

3.1 Differentially Expressed Genes

Based on the gene expression profile dataset GSE62932, a total of 479, 313, 349, and 383 DEGs with $FDR < 0.01$ and $|log2FC| \geqslant 1.5$ were identified between control samples and CRC stage I, II, III, and IV, respectively. The corresponding DEG sets were defined as DEG-stage1, DEG-stage2, DEG-stage3, and DEG-stage4.

3.2 Protein-Protein Interactions (PPIs) Networks

Four PPI networks was established based on PPIs of four DEG sets respectively (see Figure 1 for details). The network of DEG-stage1 was consisted of 4271 interaction among 424 genes, the network of DEG-stage2 was consisted of 592 interaction among 242 genes, the network of DEG-stage3 was consisted of 1040 interaction among 284 genes, and the network of DEG-stage4 was consisted of 576 interaction among 298 genes.

3.3 Functional Module Networks of Four Stages

A total of 75, 53, 57, and 67 modules were clustered by ClusterOne plugin in Cytoscape software among four PPI networks obtained above, respectively. Four function module network was constructed by functional modules identified from four PPI networks, respectively (see Figure 2 for details). Larger nodes indicate more genes in the module and the thicker edge manifest the closer the relationship between the two modules.

3.4 Functional Evolution Network of CRC

A functional evolution network was constructed among all functional modules obtained above to find modules that are closely related between CRC stage for analysis, as these modules may be closely associated with the evolution process of CRC. There were two significant evolution paths in the functional evolution network (see Figure 3 for details): the first one is a2-b1-c2- d1(Mod1), which includes 36 DEGs in total; the other is a1-b2-c1-d2(Mod2), which includes 109 DEGs in total.

3.5 Functional Enrichment Analyses

The DEGs of Mod1 were enriched in 14 pathways (details shown in Figure 4), which were strongly related to the functions of Neuroactive ligand-receptor interaction, Cytokine-cytokine receptor interaction and Chemokine signaling pathway. The DEGs of Mod2 were enriched in 8 pathways (details shown in Figure 5), and Cell cycle was the most significantly over-represented pathway among them. A total 22 pathways were identified, which were dynamically changed with pathological stages of CRC.

GO enrichment analysis was also implemented for Mod1 and Mod2 to validate above results. In addition, a total of 32 and 72 significantly over-represented GO terms were respectively obtained with $FDR < 0.05$ as the criteria. The details of top 10 of GO terms in Mod1 and Mod2 were shown in Figure 6 and Figure 7), respectively. Terms of Mod1 associated G-protein coupled receptor signaling pathway, chemokine-mediated signaling pathway, immune response and immune response were enriched among the DEGs of Mod1. Terms of Mod2 associated with cell division, mitotic nuclear division, chromosome segregation and DNA replication were enriched among the DEGs of Mod2.

A comprehensive analysis of KEGG and GO results shows that DEGs of Mod1 mainly associated with following groups of pathways: (1)signaling molecules and interaction (eg. Neuroactive ligand-receptor interaction, Cytokine-cytokine receptor interaction); (2)immune system (e.g. Chemokine signaling pathway); (3)signal transduction (eg. Calcium signaling pathway, cGMP-PKG signaling pathway); (4)human diseases (e.g. Pathways in cancer, Legionellosis, Salmonella infection). KEGG and GO results shows that DEGs of Mod2 closely related to following groups of pathways: (1)cell growth and death (e.g. Cell cycle, p53 signaling pathway, Oocyte meiosis); (2)endocrine system (e.g. Progesterone-mediated oocyte maturation); (3)replication and repair (e.g. DNA replication, Fanconi anemia pathway); (4)human diseases (e.g. Viral carcinogenesis, HTLV-I infection).

4 DISCUSSION

In this study, four DEG sets were screened between gene expression profiles of normal tissues and four stages of CRC tissues, respectively. A variety of signal pathways and GO terms closely related to functional evolution process of CRC were revealed through the function module network and the functional evolution network based on function modules of PPI networks. Function enrichment analysis indicated that DEGs involved in Mod1 and Mod2 were mainly associated with cellular processes, signal transduction information and immune system.

Figure 3 clearly visualized the fact that Mod1 and Mod2 paths that were significantly related to CRC stage development. Moreover, a total number of 22 significant KEGG pathways were identified in Mod1 and Mod2. Previous studies demonstrated that the mutation of some genes and pathways are associated with the process of CRC. Chemokine signaling affects CRC invasion and/or metastasis by altering tumor microenvironment (Itatani et al., 2016). In addition, abnormal expression of some cytokines may be related to CRC (Trost and McDonnell, 2005). Moreover, G-Protein coupled receptors (GPCRs) affect many aspects of tumorigenesis, including invasion, proliferation, motility and several cancer-related signaling pathways (Zhong and Neubig, 2001). GPCRs are pivotal regulators of several physiological processes and are corresponding targets for the treatment of cancer, as well as other tumors. Function enrichment analysis showed the DEGs in Mod2 were prominently enriched in the cell processes pathways, such as cell cycle and p53 signaling pathway. There is abundant evidence that cell cycle disruption is associated with the development of various cancers (Maddika et al., 2007; Chang et al., 2003). Numerous investigations have indicated that p53 is a tumor suppressor gene (Suzuki and Matsubara, 2011) and plays a critical role in development of tumors. It contributes to not only control cell cycle (Suzuki and Matsubara, 2011), but also keep chromosome stability (Achanta et al., 2005) and mitochondrial genetic stability (Liu et al., 2004; Iglesias et al., 2012). Consequently, results in our paper suggest that these pathways extremely likely related to CRC development, more details and evidence need to be investigated in further studies.

In conclusion, several pathways we obtained showed closely associated with CRC development in this paper. Some pathways have been proven to be involved in the development of CRC, and we predict that others might also be related to the functional evolution process of CRC, which indicates that the method of obtaining related pathways of CRC is effective. Besides, functional evolution analyses also can be performed to identify pathways in other cancer, such as liver cancer, lung cancer, which provide a new perspective to explore the dynamic evolution process of cancer.

5 ADDITIONAL REQUIREMENTS

For additional requirements for specific article types and further information please refer to Author Guidelines.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

The Author Contributions section is mandatory for all articles, including articles by sole authors. If an appropriate statement is not provided on submission, a standard one will be inserted during the production

process. The Author Contributions statement must describe the contributions of individual authors referred to by their initials and, in doing so, all authors agree to be accountable for the content of the work. Please see here for full authorship criteria.

FUNDING

Details of all funding sources should be provided, including grant numbers if applicable. Please ensure to add all necessary funding information, as after publication this is no longer possible.

ACKNOWLEDGMENTS

This work was supported by the National Natural Science Foundation of China under [Grant No. 61602386, 61972320, 61702161 and 61772426]; the Natural Science Foundation of Shaanxi Province under [Grant No. 2017JQ6008]; the Fundamental Research Funds for the Central Universities under [Grant No. 3102019DX1003]; and the Top International University Visiting Program for Outstanding Young scholars of Northwestern Polytechnical University.

SUPPLEMENTAL DATA

Supplementary Material should be uploaded separately on submission, if there are Supplementary Figures, please include the caption in the same file as the figure. LaTeX Supplementary Material templates can be found in the Frontiers LaTeX folder.

DATA AVAILABILITY STATEMENT

The datasets [GENERATED/ANALYZED] for this study can be found in the [NAME OF REPOSITORY] [LINK].

REFERENCES

Achanta, G., Sasaki, R., Feng, L., Carew, J. S., Lu, W., Pelicano, H., et al. (2005). Novel role of p53 in maintaining mitochondrial genetic stability through interaction with dna pol γ. *The EMBO journal* 24, 3482–3492

Carlson, M., Falcon, S., Pages, H., and Li, N. (2016). hgu133plus2. db: Affymetrix human genome u133 plus 2.0 array annotation data (chip hgu133plus2). *R package version* 3

Chang, F., Lee, J., Navolanic, P., Steelman, L., Shelton, J., Blalock, W., et al. (2003). Involvement of pi3k/akt pathway in cell cycle progression, apoptosis, and neoplastic transformation: a target for cancer chemotherapy. *Leukemia* 17, 590

Ferreira, J. A. (2007). The benjamini-hochberg method in the case of discrete test statistics. *The international journal of biostatistics* 3

Gautier, L., Cope, L., Bolstad, B. M., and Irizarry, R. A. (2004). affy-analysis of affymetrix genechip data at the probe level. *Bioinformatics* 20, 307–315

Gulmann, C., Sheehan, K. M., Conroy, R. M., Wulfkuhle, J. D., Espina, V., Mullarkey, M. J., et al. (2009). Quantitative cell signalling analysis reveals down-regulation of mapk pathway activation in colorectal cancer. *The Journal of Pathology: A Journal of the Pathological Society of Great Britain and Ireland* 218, 514–519

Iglesias, P., Salas, A., and Costoya, J. A. (2012). The maintenance of mitochondrial genetic stability is crucial during the oncogenic process. *Communicative & integrative biology* 5, 34–38

Irizarry, R. A., Hobbs, B., Collin, F., Beazer-Barclay, Y. D., Antonellis, K. J., Scherf, U., et al. (2003). Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 4, 249–264

Itatani, Y., Kawada, K., Inamoto, S., Yamamoto, T., Ogawa, R., Taketo, M., et al. (2016). The role of chemokines in promoting colorectal cancer invasion/metastasis. *International journal of molecular sciences* 17, 643

Klaus, A. and Birchmeier, W. (2008). Wnt signalling and its impact on development and cancer. *Nature Reviews Cancer* 8, 387

Li, X.-L., Zhou, J., Chen, Z.-R., and Chng, W.-J. (2015). P53 mutations in colorectal cancer-molecular pathogenesis and pharmacological reactivation. *World journal of gastroenterology: WJG* 21, 84

Liu, G., Parant, J. M., Lang, G., Chau, P., Chavez-Reyes, A., El-Naggar, A. K., et al. (2004). Chromosome stability, in the absence of apoptosis, is critical for suppression of tumorigenesis in trp53 mutant mice. *Nature genetics* 36, 63

Maddika, S., Ande, S. R., Panigrahi, S., Paranjothy, T., Weglarczyk, K., Zuse, A., et al. (2007). Cell survival, cell death and cell cycle pathways are interconnected: implications for cancer therapy. *Drug Resistance Updates* 10, 13–29

Nepusz, T., Yu, H., and Paccanaro, A. (2012). Detecting overlapping protein complexes in protein-protein interaction networks. *Nature methods* 9, 471

Oh, S., Gwak, J., Park, S., and Yang, C. S. (2014). Green tea polyphenol egcg suppresses w nt/β-catenin signaling by promoting gsk-3β-and pp2a-independent β-catenin phosphorylation/degradation. *Biofactors* 40, 586–595

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W., et al. (2015). limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic acids research* 43, e47–e47

Russo, A., Bazan, V., Iacopetta, B., Kerr, D., Soussi, T., and Gebbia, N. (2005). The tp53 colorectal cancer international collaborative study on the prognostic and predictive significance of p53 mutation: influence of tumor site, type of mutation, and adjuvant treatment. *Journal of clinical oncology* 23, 7518–7528

Rychahou, P. G., Murillo, C. A., and Evers, B. M. (2005). Targeted rna interference of pi3k pathway components sensitizes colon cancer cells to tnf-related apoptosis-inducing ligand (trail). *Surgery* 138, 391–397

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research* 13, 2498–2504

Sherman, B. T., Lempicki, R. A., et al. (2009). Systematic and integrative analysis of large gene lists using david bioinformatics resources. *Nature protocols* 4, 44–57

Shim, A.-R., Dong, G.-z., Lee, H. J., and Ryu, J.-H. (2015). Atractylochromene is a repressor of wnt/β-catenin signaling in colon cancer cells. *Biomolecules & therapeutics* 23, 26

Siegel, R. L., Miller, K. D., and Jemal, A. (2019). Cancer statistics, 2019. *CA: a cancer journal for clinicians* 69, 7–34

Smyth, G. K., Ritchie, M., Thorne, N., and Wettenhall, J. (2005). Limma: linear models for microarray data. in bioinformatics and computational biology solutions using r and bioconductor. statistics for biology and health

Suzuki, K. and Matsubara, H. (2011). Recent advances in p53 research and cancer treatment. *BioMed Research International* 2011

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., et al. (2014). String v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic acids research* 43, D447–D452

Trost, L. and McDonnell, J. (2005). Important cutaneous manifestations of inflammatory bowel disease. *Postgraduate medical journal* 81, 580–585

Xiong, B., Gong, L.-L., Zhang, F., Hu, M.-B., and Yuan, H.-Y. (2002). Tgf-β1 expression and angiogenesis in colorectal cancer tissue. *World journal of gastroenterology* 8, 496

You, S., Zhou, J., Chen, S., Zhou, P., Lv, J., Han, X., et al. (2010). Ptch1, a receptor of hedgehog signaling pathway, is correlated with metastatic potential of colorectal cancer. *Upsala journal of medical sciences* 115, 169–175

Zhong, H. and Neubig, R. R. (2001). Regulator of g protein signaling proteins: novel multifunctional drug targets. *Journal of Pharmacology and Experimental Therapeutics* 297, 837–845

FIGURE CAPTIONS

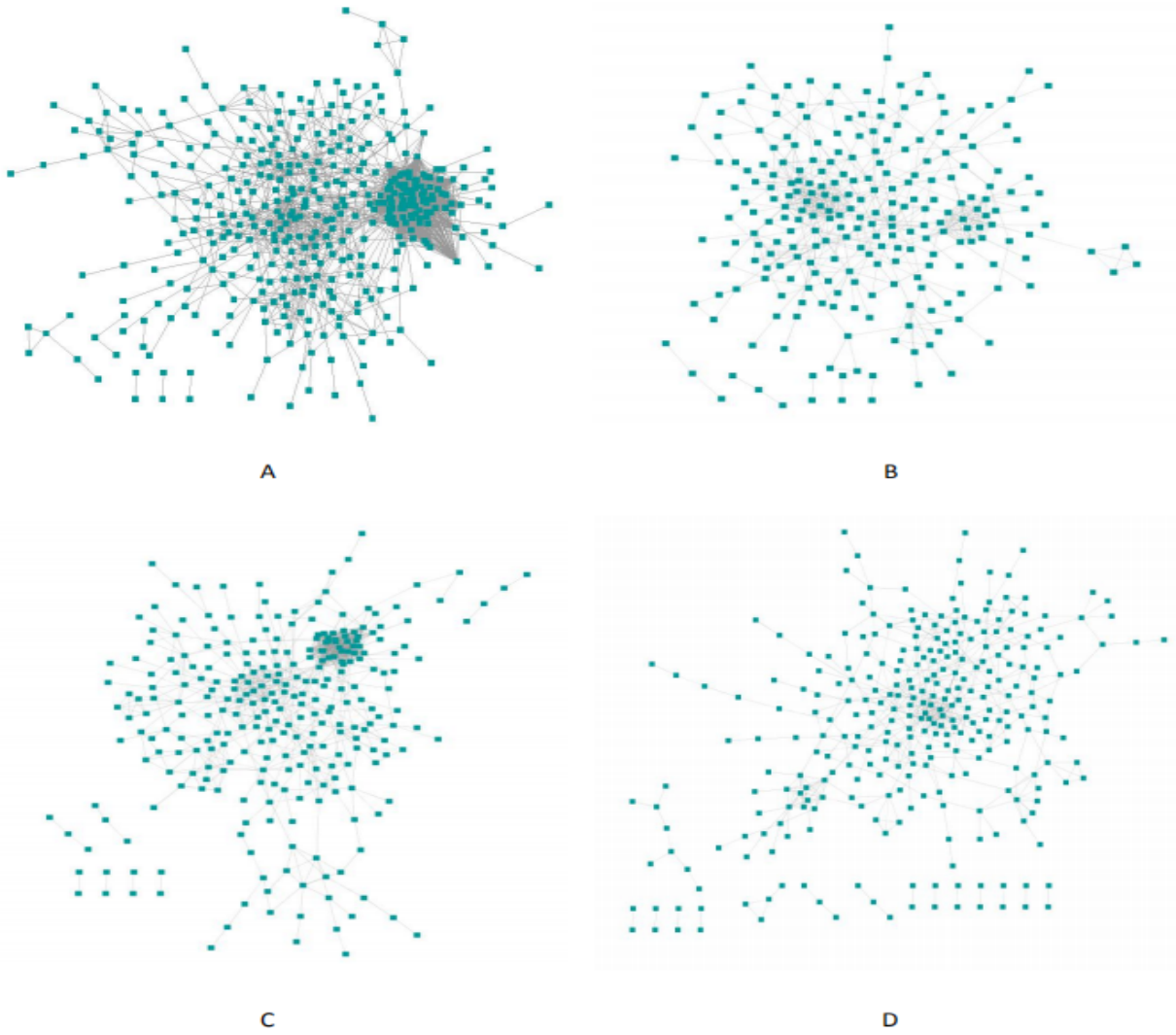

Figure 1. PPI networks of four stages of CRC: vertexs represent the DEGs; edges represent the interactions among genes.**(A)** is PPI network of DEG-stage1, **(B)** is PPI network of DEG-stage2, **(C)** is PPI network of DEG-stage3, **(D)** is PPI network of DEG-stage4.

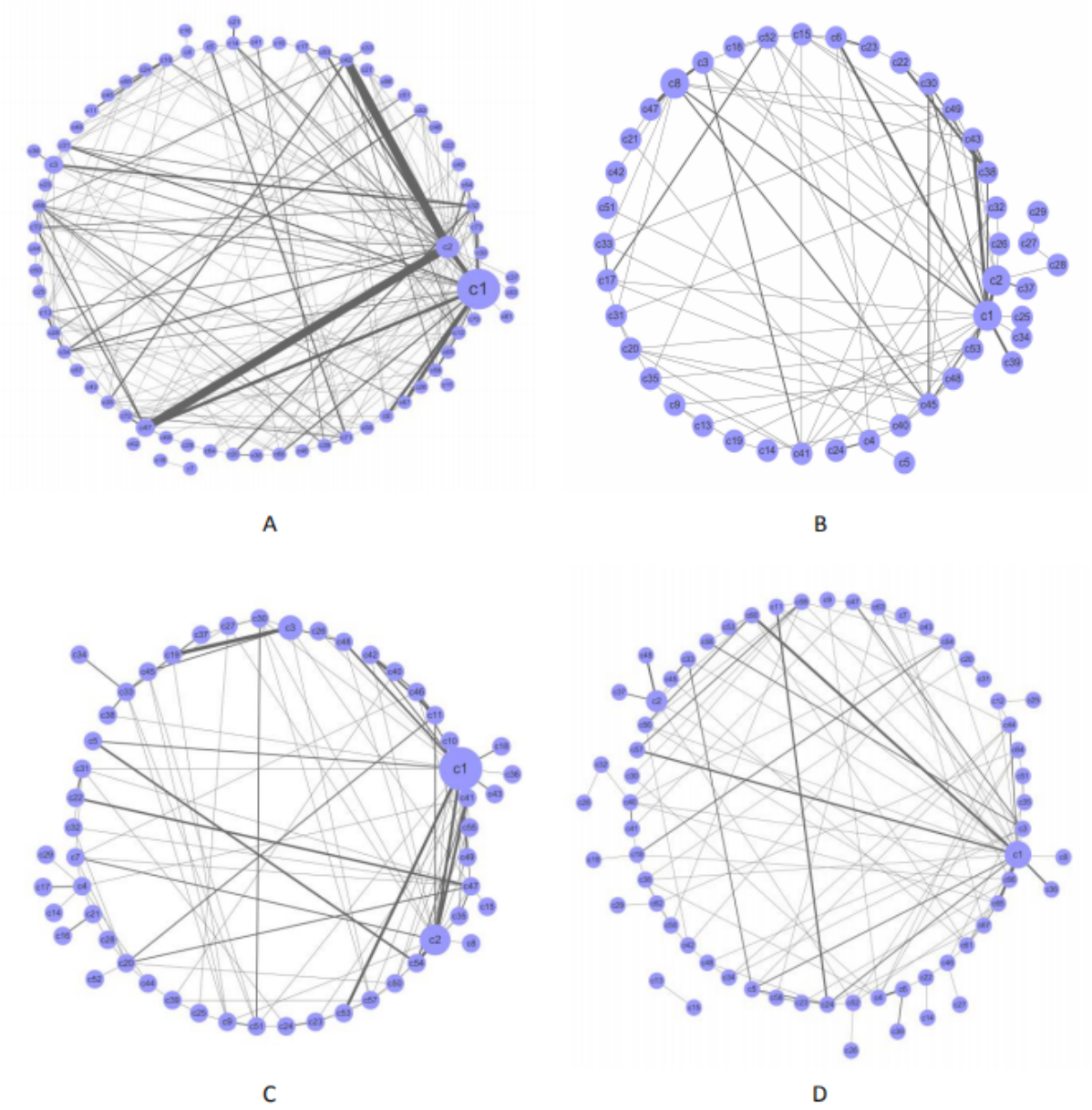

Figure 2. Function module networks of four stages of CRC: vertexs represent the function modules; edges represent the relationship among modules. **(A)** is function module network of DEG-stage1, **(B)** function module network of DEG-stage2, **(C)** function module network of DEG-stage3, **(D)** function module network of DEG-stage4.

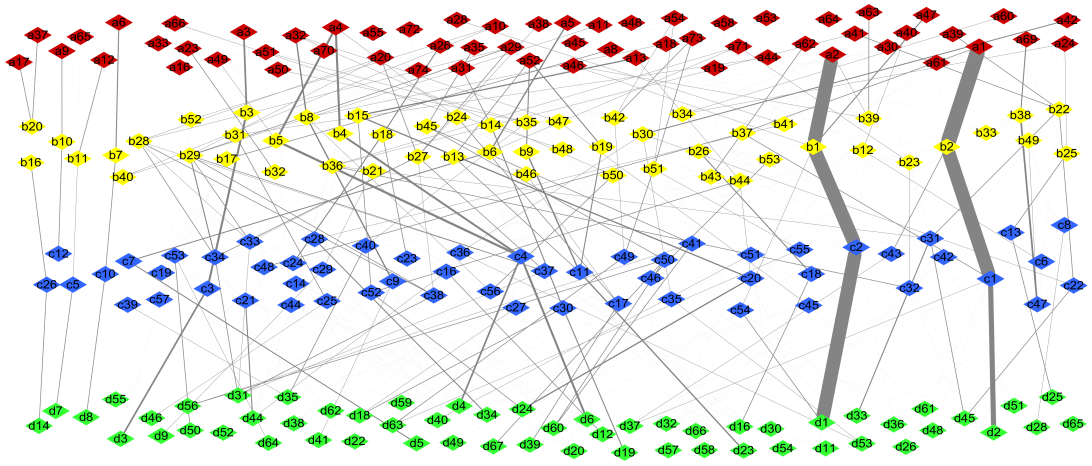

Figure 3. Functional evolution networkbetween adjacent stages of CRC: vertexs represent the function modules; edges represent the relationship among modules. Red, yellow, blue, and green circles represent functional modules belonging to DEG-stage1, DEG-stage2, DEG-stage3, and DEG-stage4, respectively.

Category	Term	Count	PValue	Genes
KEGG_PATHWAY	hsa04062:Chemokine signaling pathway	12	3.76E-11	CXCL1, CXCL3, CXCL16, CCR10, CXCL2, CCL8, CXCL8, GNG2, CXCL11, CXCL12, CCL28, GNG7
KEGG_PATHWAY	hsa04060:Cytokine-cytokine receptor interaction	12	6.75E-10	CXCL1, IL1R2, CXCL3, CXCL16, CCR10, CXCL2, CCL8, CXCL8, IL6R, CXCL11, CXCL12, CCL28
KEGG_PATHWAY	hsa04080:Neuroactive ligand-receptor interaction	10	6.05E-07	EDNRB, AGTR1, ADRB2, ADRB1, P2RY14, OXTR, LPAR1, BDKRB2, NPY1R, VIPR1
KEGG_PATHWAY	hsa04020:Calcium signaling pathway	7	5.36E-05	EDNRB, AGTR1, ADRB2, ADRB1, GNA11, OXTR, BDKRB2
KEGG_PATHWAY	hsa05200:Pathways in cancer	9	8.96E-05	EDNRB, AGTR1, GNA11, CXCL8, GNG2, LPAR1, BDKRB2, CXCL12, GNG7
KEGG_PATHWAY	hsa04022:cGMP-PKG signaling pathway	6	3.22E-04	EDNRB, AGTR1, ADRB2, ADRB1, GNA11, BDKRB2
KEGG_PATHWAY	hsa05134:Legionellosis	4	0.001170696	CXCL1, CXCL3, CXCL2, CXCL8
KEGG_PATHWAY	hsa05132:Salmonella infection	4	0.00402044	CXCL1, CXCL3, CXCL2, CXCL8
KEGG_PATHWAY	hsa04672:Intestinal immune network for IgA production	3	0.014386716	CCR10, CXCL12, CCL28
KEGG_PATHWAY	hsa04923:Regulation of lipolysis in adipocytes	3	0.02005701	ADRB2, ADRB1, NPY1R
KEGG_PATHWAY	hsa04621:NOD-like receptor signaling pathway	3	0.02005701	CXCL1, CXCL2, CXCL8
KEGG_PATHWAY	hsa04924:Renin secretion	3	0.025757512	AGTR1, ADRB2, ADRB1
KEGG_PATHWAY	hsa04024:cAMP signaling pathway	4	0.041367645	ADRB2, ADRB1, OXTR, NPY1R
KEGG_PATHWAY	hsa04540:Gap junction	3	0.046182301	ADRB1, GNA11, LPAR1

Figure 4. KEGG pathways of a2-b1-c2-d1(Mod1)

Category	Term	Count	PValue	Genes
KEGG_PATHWAY	hsa04110:Cell cycle	15	2.99E-16	CDC6, CDK1, RBL1, TTK, CHEK1, CDC20, PTTG1, MCM2, CDC25B, CCNB1, CCNB2, MAD2L1, CDKN2B, BUB1, CCNA2
KEGG_PATHWAY	hsa04114:Oocyte meiosis	8	1.09E-06	CCNB1, CDK1, MAD2L1, CCNB2, BUB1, AURKA, CDC20, PTTG1
KEGG_PATHWAY	hsa04914:Progesterone-mediated oocyte maturation	7	4.16E-06	CCNB1, CDK1, MAD2L1, CCNB2, BUB1, CCNA2, CDC25B
KEGG_PATHWAY	hsa04115:p53 signaling pathway	5	3.43E-04	CCNB1, CDK1, CCNB2, RRM2, CHEK1
KEGG_PATHWAY	hsa05203:Viral carcinogenesis	6	0.003507217	CDK1, CDKN2B, RBL1, CHEK1, CDC20, CCNA2
KEGG_PATHWAY	hsa03030:DNA replication	3	0.014217362	RFC3, MCM2, RNASEH2A
KEGG_PATHWAY	hsa03460:Fanconi anemia pathway	3	0.029469383	FANCI, BRCA2, UBE2T
KEGG_PATHWAY	hsa05166:HTLV-I infection	5	0.038742447	MAD2L1, CDKN2B, CHEK1, CDC20, PTTG1

Figure 5. KEGG pathways of a1-b2-c1-d2(Mod2)

Category	Term	Count	PValue	Genes
GOTERM_BP_DIRECT	GO:0007186~G-protein coupled receptor signaling pathway	23	6.16E-20	VIP, CXCL1, FFAR4, GNA11, CXCL3, CXCL2, ACKR1, CXCL8, CCL8, OXTR, LPAR1,
GOTERM_BP_DIRECT	GO:0070098~chemokine-mediated signaling pathway	9	1.46E-12	CXCL1, CXCL3, CCR10, CXCL2, ACKR1, CCL8, CXCL8, CXCL11, CXCL12
GOTERM_BP_DIRECT	GO:0006935~chemotaxis	9	1.23E-10	CXCL1, CXCL16, CCR10, CXCL2, CCL8, CXCL8, CXCL11, CXCL12, CCL28
GOTERM_BP_DIRECT	GO:0060326~cell chemotaxis	7	3.95E-09	CXCL1, AGTR1, CXCL2, CCL8, LPAR1, CXCL12, CCL28
GOTERM_BP_DIRECT	GO:0006955~immune response	11	9.25E-09	CXCL1, IL1R2, CXCL3, CCR10, CXCL2, CCL8, CXCL8, CXCL11, VIPR1, CXCL12, CCL28
GOTERM_BP_DIRECT	GO:0007204~positive regulation of cytosolic calcium ion concentration	8	9.82E-09	EDNRB, AGTR1, CCR10, TAC1, OXTR, LPAR1, BDKRB2, CCL28
GOTERM_BP_DIRECT	GO:0006954~inflammatory response	10	5.81E-08	CXCL1, CXCL3, CXCL2, ACKR1, CCL8, CXCL8, TAC1, BDKRB2, CXCL11, CXCL12
GOTERM_BP_DIRECT	GO:0007267~cell-cell signaling	8	7.81E-07	EDN3, ADRB2, ADRB1, CCL8, TAC1, PYY, CXCL11, SST
GOTERM_BP_DIRECT	GO:0007166~cell surface receptor signaling pathway	8	1.30E-06	EDNRB, EDN3, ADRB2, OXTR, BDKRB2, NPY1R, VIPR1, SST
GOTERM_BP_DIRECT	GO:0090023~positive regulation of neutrophil chemotaxis	4	1.24E-05	CXCL1, CXCL3, CXCL2, CXCL8

Figure 6. Top 10 of GO(BP) term of a2-b1-c2-d1(Mod1)

Category	Term	Count	PValue	Genes
GOTERM_BP_DIRECT	GO:0051301~cell division	39	1.16E-37	HAUS6, NEK2, AURKA, PTTG1, SPC25, KIF2C, NCAPH, NCAPG2, NCAPG, OIP5, CDCA2, BUB1, SKA3, CCNA2, CDCA5, HELLS, CDCA3, KIF14, CDK1, CDC6, KIF11, PSRC1, NUF2, TPX2, KIF18B, NDC80, SPDL1, CENPE, CDC20, BIRC5, UBE2C, SMC2, CDC25B, CCNB1, CCNB2, MAD2L1, CKS2, CENPW, UBE2S
GOTERM_BP_DIRECT	GO:0007067~mitotic nuclear division	32	1.98E-32	HAUS6, NEK2, ANLN, AURKA, CEP55, AURKB, PTTG1, SPC25, KIF2C, NCAPG2, OIP5, BUB1, CDCA2, SKA3, CCNA2, CDCA5, HELLS, ASPM, CDCA3, CDK1, CENPN, CDC6, KIF11, NUF2, TPX2, NDC80, BIRC5, CDC20, PBK, CDC25B, CCNB2, CENPW
GOTERM_BP_DIRECT	GO:0007059~chromosome segregation	13	8.08E-15	CENPN, KIF11, NEK2, NUF2, NDC80, CENPE, SPC25, HJURP, OIP5, CDCA2, SKA3, CENPW, TOP2A
GOTERM_BP_DIRECT	GO:0007062~sister chromatid cohesion	14	5.42E-14	CENPN, KIF18A, NUF2, CDC20, CENPE, SPDL1, BIRC5, NDC80, AURKB, KIF2C, SPC25, MAD2L1, BUB1, CDCA5
GOTERM_BP_DIRECT	GO:0000086~G2/M transition of mitotic cell cycle	15	1.09E-13	CDK1, HAUS6, NEK2, FOXM1, TPX2, AURKA, BIRC5, CHEK1, CDC25B, HMMR, AJUBA, CCNB1, CCNB2, CDKN2B, MELK
GOTERM_BP_DIRECT	GO:0007080~mitotic metaphase plate congression	9	5.07E-11	KIF14, CCNB1, KIF2C, PSRC1, KIF18A, CENPE, SPDL1, CEP55, CDCA5
GOTERM_BP_DIRECT	GO:0007018~microtubule-based movement	10	1.56E-09	KIF23, KIF14, KIF2C, KIF4A, KIF11, KIF18A, KIF18B, CENPE, KIF26B, KIF20A
GOTERM_BP_DIRECT	GO:0006260~DNA replication	12	2.79E-09	EXO1, CDK1, CDC6, RFC3, MCM8, DTL, RRM2, CHEK1, MCM2, RNASEH2A, MCM10, DSCC1
GOTERM_BP_DIRECT	GO:0007052~mitotic spindle organization	7	2.42E-08	CCNB1, STIL, SPC25, KIF11, TTK, AURKA, NDC80
GOTERM_BP_DIRECT	GO:0031145~anaphase-promoting complex-dependent catabolic process	9	2.77E-08	CCNB1, CDK1, MAD2L1, AURKA, CDC20, PTTG1, AURKB, UBE2C, UBE2S
GOTERM_BP_DIRECT	GO:0007051~spindle organization	6	3.26E-08	KIF11, TTK, AURKA, AURKB, ASPM, AUNIP

Figure 7. Top 10 of GO(BP) term of a1-b2-c1-d2(Mod2)