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

Bolin Chen[†‡], Manting Yang[†], Li Gao[†] and Xuequn Shang[†‡*]

[†] School of Computer Science, Northwestern Polytechnical University, Xi'an, China
[‡]Key Laboratory of Big Data Storage and Management, Ministry of Industry and Information Technology, Northwestern Polytechnical University, Xi'an, China
[*]Corresponding email: npu_bioinf@hotmail.com

***Abstract*—This paper presents a new method for identifying pathways associated with colorectal cancer (CRC) development with bioinformatics technology. We obtained CRC samples and healthy sample from the database of GSE62932, and limma algorithm was used to identify four differentially expressed genes (DEGs) sets between CRC stage I, II, III, and IV samples and control samples. KEGG pathways enrichment analysis was performed for four DEG sets with the ClueGo tool. Analyzing the relationships of those pathways and visualizing an interactive network of pathways with Cytoscape. Protein-protein interactions (PPIs) of four DEG sets were analyzed with STRING, and four networks were visualized and functional modules were screened out, respectively. A total of 479, 313, 349, and 383 DEGs were identified and they were enriched in 17, 16, 20, and 24 pathways, respectively. An interactive network was constructed within 35 pathways. A total of 75, 53, 57, and 67 modules were identified among four PPI networks, respectively, and a network of the modules between adjacent phases was constructed. There were 2 significant paths in the network: one is a2-b1-c2-d1(Mod1) enriched in 10 pathways among 36 DEGs totally, the other is a1-b2-c1-d2(Mod2) enriched in 7 pathways among 109 DEGs totally. The significant KEGG pathways obtained in present study may play critical roles in the development of CRC, which contributes to explore molecular mechanisms of CRC development.**

***Index Terms*—colorectal cancer, differentially expressed gene, dynamic network analysis, functional evolution process**

I. 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 [1]. Therefore, exploring the molecular mechanism of CRC is conducive to the detection of early CRC, and promotes the development of new screening methods, prognostic markers and therapeutic methods.

Numerous studies have shown that CRC is a complex disease with multiple genes, multiple stages, and multiple factors. The occurrence of CRC is closely related to the mutation of the cell signaling pathway. In recent years, several cell signal transduction pathways related to CRC carcinogenesis have been discovered, including Wnt/-catenin [2]–[4], Hedgehog [5], TGF-/Smads [6], PI3K/Akt [7], MAPK [8] and p53 [9], [10] 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.

The common biological method for studying cancer is to screen differentially expressed genes between control tissues and cancer tissues, and perform GO enrichment analysis, KEGG pathway analysis, protein interaction (PPI) network analysis and module analysis on significant differential genes, respectively. To explore the molecular mechanisms of cancer by identifying and analyzing key genes and related pathways of cancer.

The various signal pathways in the cell are not isolated, but multiple signal transduction pathways interact and correlate to form a complex signal network system [11]. Therefore, analysis of the interactions between pathways at multiple stages of cancer may be more helpful in analyzing the molecular mechanisms of cancer development.

In this paper, we propose a new approach to study the molecular mechanisms of the CRC development. Firstly, we screened the differentially expressed genes (DEGs) between CRC stage I, II, III, and IV samples and healthy controls based on the microarray data, respectively. Next, pathway enrichment analysis and PPI network module analysis were performed on four sets DEGs, respectively. Finally, construction of pathway networks of four sets DEGs and integration analysis of four PPI networks and modules were contributed to explore the critical signal pathways related to CRC development.

II. Materials and Methods

A. Data Sources

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.

B. Data Preprocessing

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 [12], [13]. After filtering out nonspecific probes, the rest probe level data were mapped to corresponding gene symbols using the annotation information package in GPL570, hgu133plus2.db [14]. 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.

C. 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 [15], [16] algorithm. Using the Benjamini-Hochberg [17] method to correct p-value to acquire FDR. Calculating Fold Change (FC) value of each gene between two groups. Only the genes that meets $FDR < 0.01$ and $|log2FC| \geqslant 1.5$ were defined as DEGs.

D. Pathway Enrichment Analysis

The ClueGO [18] and CluePedia [19], two plugins of the Cytoscape software [20], were used to conduct pathway enrichment analysis to identify the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of DEGs and visualize relationships of those pathways. In this study, the over-represented KEGG pathways were investigated through enrichment/depletion (Twosided hypergeometric test) and Benjamini-Hochberg correction methods (criteria: $pV \leqslant 0.05$, Kappa Score threshold=0.4).

E. Interaction Network of Pathways at All Stages

The ClueGo [18] plugin was used to visualize the relationships between the significantly enriched pathways of DEGs at various stages. Then analyze the relationship of all pathways based on Kappa Score. An interaction pathway network was constructed and visualized by using Cytoscape software [20]. Each pathway in the network comes from one or more groups of DEGs, which was colored to identify the stages. If a pathway appears more than one stage, we will divide the node into the same number parts and use different colors to represent these stages. A higher degree of pathway indicates more important role in the signaling network.

F. Protein-Protein Interactions (PPIs) of DEGs

STRING (Search Tool for the Retrieval of Interacting Genes) [21] 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 [20].

G. Module Analysis between Adjacent Stages

ClusterONE [22], a Cytoscape plugin, was employed to identify functional modules (criteria: Min size=3, overlap threshold=0.8, simalarity = coefficient) in four stages of PPI network, respectively. Since there are common differential genes in different stages, the modules of different stages of the PPI network also have the same DEGs, which makes it possible to calculate overlapping score between modules of adjacent phases. A module relationship network was then generated and it was visualized in Cytoscape [20]. Closely related modules among the four stages of CRC were then identified according to the network. Moreover, ClueGo [18] was used to identify the KEGG pathways (criteria: $p-value \leqslant 0.05$, Kappa Score threshold=0.4) involving the DEGs in four interaction modules in four stages of CRC.

III. Results

A. 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, and the corresponding DEG sets were defined as DEG-stage1, DEG-stage2, DEG-stage3, and DEG-stage4.

B. Pathway Enrichment Analysis of four DEG sets

KEGG pathway enrichment analysis was performed, and four DEG sets were significantly enriched in 17, 16, 20, and 24 pathways. The details of them can be found in Table I, Table II, Table III and table IV, respectively.

TABLE I
KEGG PATHWAYS OF DEG-STAGE1

ID	Term	*p-value*	Count
KEGG:04514	Cell adhesion molecules (CAMs)	5.14E-03	10
KEGG:04924	Renin secretion	2.34E-03	7
KEGG:05146	Amoebiasis	4.04E-03	8
KEGG:00561	Glycerolipid metabolism	1.13E-03	7
KEGG:00910	Nitrogen metabolism	9.73E-03	3
KEGG:04115	p53 signaling pathway	2.99E-03	7
KEGG:04512	ECM-receptor interaction	6.17E-03	7
KEGG:05222	Small cell lung cancer	1.20E-02	7
KEGG:04110	Cell cycle	9.37E-07	15
KEGG:04114	Oocyte meiosis	6.19E-03	9
KEGG:04914	Progesterone-mediated oocyte maturation	1.27E-03	9
KEGG:00040	Pentose and glucuronate interconversions	2.45E-04	6
KEGG:00053	Ascorbate and aldarate metabolism	5.39E-03	4
KEGG:00140	Steroid hormone biosynthesis	5.13E-03	6
KEGG:00830	Retinol metabolism	8.76E-03	6
KEGG:00982	Drug metabolism	2.99E-03	7
KEGG:05204	Chemical carcinogenesis	1.49E-03	8

The total number of non-overlapped pathways was 35 and four DEG sets were commonly enriched in the following 8 categories, which were (1) Cell adhesion molecules (CAMs), (2) Nitrogen metabolism, (3) Pentose and glucuronate interconversions, (4) Ascorbate and aldarate metabolism, (5) Steroid hormone biosynthesis, (6) Retinol metabolism, (7) Drug metabolism, and (8) Chemical carcinogenesis.

To be more specific, DEG-stage1 (Table I) was mainly enriched in Cell Cycle; DEG-stage2 (Table II) was mainly

TABLE II
KEGG PATHWAYS OF DEG-STAGE2

ID	Term	*p-value*	Count
KEGG:04514	Cell adhesion molecules (CAMs)	4.49E-03	8
KEGG:04978	Mineral absorption	1.33E-02	4
KEGG:00480	Glutathione metabolism	1.83E-02	4
KEGG:04060	Cytokine-cytokine receptor interaction	1.10E-02	12
KEGG:04672	Intestinal immune network for IgA production	1.16E-02	4
KEGG:00040	Pentose and glucuronate interconversions	3.30E-04	5
KEGG:00053	Ascorbate and aldarate metabolism	1.24E-02	3
KEGG:00140	Steroid hormone biosynthesis	2.29E-02	4
KEGG:00910	Nitrogen metabolism	2.04E-04	4
KEGG:04964	Proximal tubule bicarbonate reclamation	7.89E-03	3
KEGG:04976	Bile secretion	9.59E-03	5
KEGG:00071	Fatty acid degradation	7.97E-03	4
KEGG:00830	Retinol metabolism	1.92E-04	7
KEGG:00980	Metabolism of xenobiotics by cytochrome P450	2.44E-03	6
KEGG:00982	Drug metabolism	1.85E-03	6
KEGG:05204	Chemical carcinogenesis	6.71E-04	7

TABLE III
KEGG PATHWAYS OF DEG-STAGE3

ID	Term	*p-value*	Count
KEGG:04024	cAMP signaling pathway	1.38E-02	10
KEGG:04514	Cell adhesion molecules (CAMs)	1.08E-02	8
KEGG:04978	Mineral absorption	4.16E-03	5
KEGG:00230	Purine metabolism	1.95E-02	7
KEGG:00910	Nitrogen metabolism	1.95E-05	5
KEGG:04964	Proximal tubule bicarbonate reclamation	1.19E-02	3
KEGG:04976	Bile secretion	3.92E-03	6
KEGG:00053	Ascorbate and aldarate metabolism	2.27E-03	4
KEGG:00232	Caffeine metabolism	8.85E-05	3
KEGG:00983	Drug metabolism	1.30E-03	7
KEGG:05204	Chemical carcinogenesis	7.42E-06	10
KEGG:00040	Pentose and glucuronate interconversions	6.50E-05	6
KEGG:00053	Ascorbate and aldarate metabolism	2.27E-03	4
KEGG:00071	Fatty acid degradation	1.33E-02	4
KEGG:00140	Steroid hormone biosynthesis	1.54E-03	6
KEGG:00830	Retinol metabolism	7.30E-05	8
KEGG:00980	Metabolism of xenobiotics by cytochrome P450	5.12E-03	6
KEGG:00982	Drug metabolism	3.92E-03	6
KEGG:05204	Chemical carcinogenesis	7.42E-06	10

TABLE IV
KEGG PATHWAYS OF DEG-STAGE4

ID	Term	*p-value*	Count
KEGG:04060	Cytokine-cytokine receptor interaction	1.62E-02	12
KEGG:04072	Phospholipase D signaling pathway	9.04E-03	8
KEGG:00564	Glycerophospholipid metabolism	1.24E-02	6
KEGG:02010	ABC transporters	1.18E-02	4
KEGG:04215	Apoptosis	2.70E-02	3
KEGG:04261	Adrenergic signaling in cardiomyocytes	2.50E-02	7
KEGG:04514	Cell adhesion molecules (CAMs)	7.71E-03	8
KEGG:04540	Gap junction	7.85E-03	6
KEGG:04024	cAMP signaling pathway	3.19E-03	11
KEGG:04924	Renin secretion	1.17E-02v 5	
KEGG:00910	Nitrogen metabolism	1.47E-05	5
KEGG:04964	Proximal tubule bicarbonate reclamation	1.01E-02	3
KEGG:04976	Bile secretion	2.93E-03	6
KEGG:04151	PI3K-Akt signaling pathway	5.01E-03	15
KEGG:04512	ECM-receptor interaction	2.31E-02	5
KEGG:05146	Amoebiasis	6.03E-04	8
KEGG:05222	Small cell lung cancer	1.02E-02	6
KEGG:00040	Pentose and glucuronate interconversions	5.02E-04	5
KEGG:00053	Ascorbate and aldarate metabolism	1.83E-03	4
KEGG:00140	Steroid hormone biosynthesis	6.53E-03	5
KEGG:00830	Retinol metabolism	2.03E-03	6
KEGG:00980	Metabolism of xenobiotics by cytochrome P450	3.84E-03	6
KEGG:00982	Drug metabolism	5.26E-04	7
KEGG:05204	Chemical carcinogenesis	5.59E-03	6

enriched in Nitrogen metabolism and Retinol metabolism; DEG-stage3 (Table III) was mainly enriched in Chemical carcinogenesis, and DEG-stage4 (Table IV) was mainly enriched in Nitrogen metabolism and Drug metabolism.

C. Interaction Network of Pathways at four Stages of CRC

In order to explore the development process of various stages of CRC, a network of pathways was constructed by taking the above 35 pathways as vertices and their functional relationship as edges. The details of this network was illustrated in Fig 1). The relationships between pathways was evaluated according to the Kappa Scores, which contained the following two types: (a) one was pathway-pathway, and (b) the other was pathway-gene.

Each vertex was represented in a color pie chat, where red, yellow, blue, and green represent pathways part of DEG-stage1, DEG-stage2, DEG-stage3, and DEG-stage4, respectively. Significantly interaction pathways were marked with thick lines in the pathway network, which were considered to plays a critical effect in the development of cancer.

In this network, there were 8 pathways connected closely each other, including (1) Pentose and glucuronate interconversions, (2) Ascorbate and aldarate metabolism, (3) Steroid hormone biosynthesis, (4) Retinol metabolism, (5) Drug

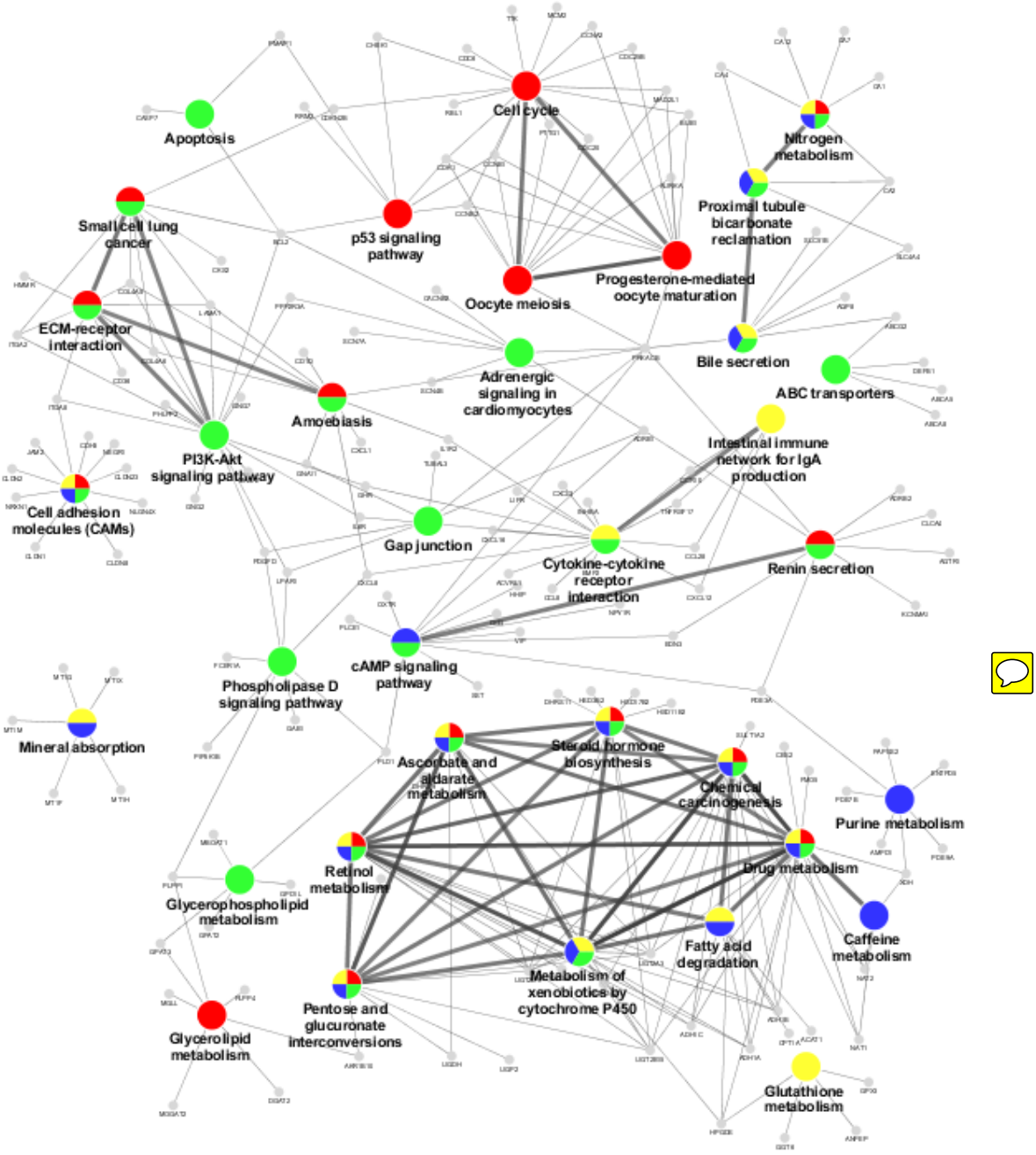
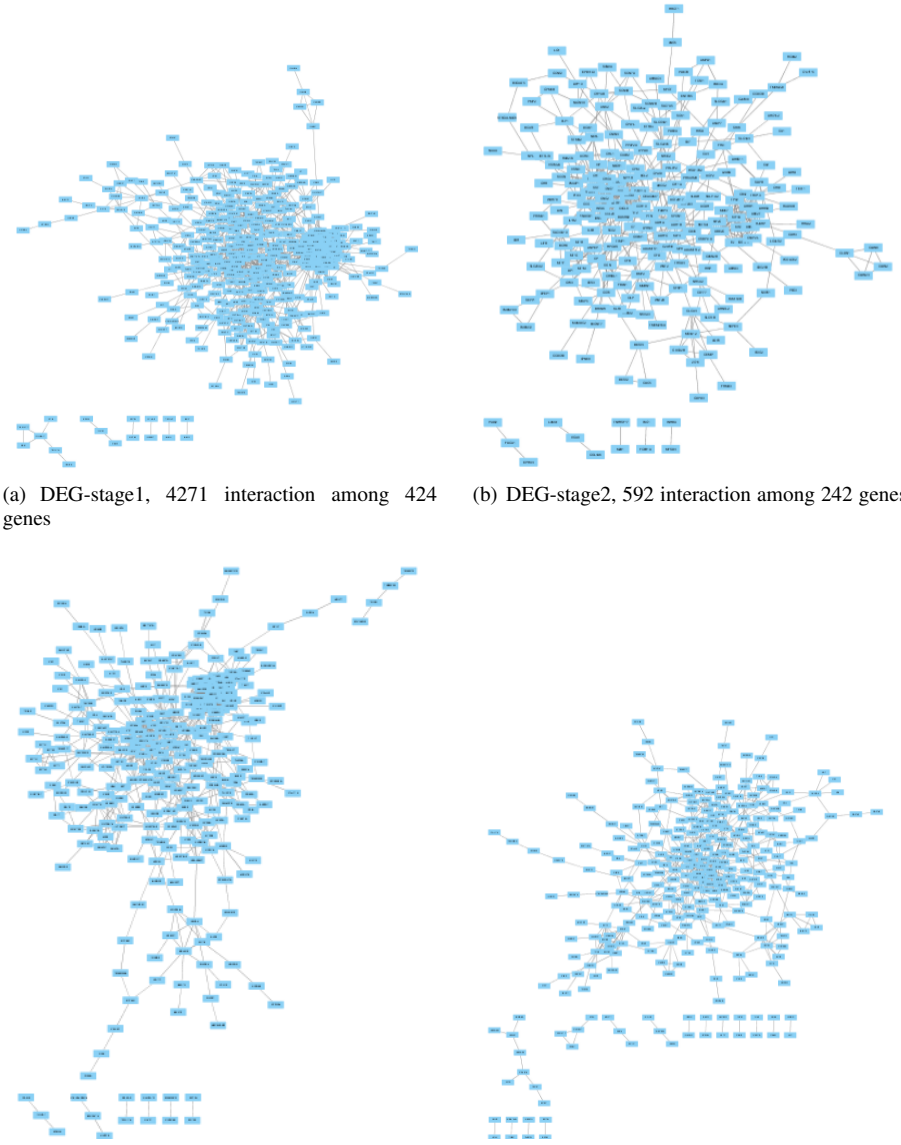


Fig. 1. Interaction network of pathways at four stages of CRC. Red, yellow, blue, and green represent pathways part of DEG-stage1, DEG-stage2, DEG-stage3, and DEG-stage4, respectively.



(a) DEG-stage1, 4271 interaction among 424 genes

(b) DEG-stage2, 592 interaction among 242 genes

(c) DEG-stage3, 1040 interaction among 284 genes

(d) DEG-stage4, 576 interaction among 298 genes

Fig. 2. PPI networks of four stage of CRC

metabolism, (6) Chemical carcinogenesis, (7) Metabolism of xenobiotics by cytochrome P450, (8) Fatty acid degradation.

D. Protein-Protein Interactions (PPIs) of four DEG sets

Four PPI networks was established based on PPIs of four DEG sets respectively (see Fig 2 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.

E. Modules Analysis between Adjacent Stages of CRC

A total of 75, 53, 57, and 67 modules were identified among four PPI networks obtained above, respectively. The relationships of modules between adjacent stages was assessed by the number of common genes in both modules. There were two obvious paths in the network (see Fig 3 for details): the first one is a2-b1-c2-d1(Mod1), which includes 36 DEGs in total; the other one is a1-b2-c1-d2(Mod2), which includes 109 DEGs in total. The DEGs in the former group were enriched in 10 pathways (details shown in Table V), which were strongly related to the functions of Neuroactive ligand-receptor interaction and Chemokine signaling pathway. The DEGs in the latter group was enriched in 7 pathways (details shown in Table VI), which were strongly related to the function of

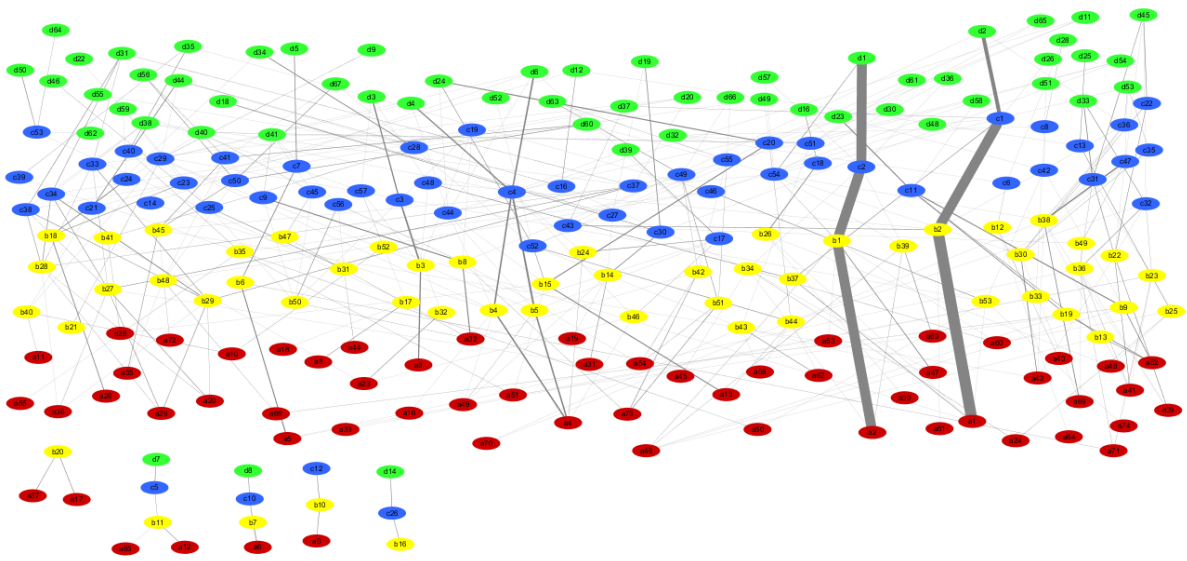


Fig. 3. Modules relation network between adjacent stages of CRC. Red, yellow, blue, and green circles represent modules belonging to DEG-stage1, DEG-stage2, DEG-stage3, and DEG-stage4, respectively.

Cell Cycle. A total 17 pathways were identified, which were dynamically changed with pathological stages of CRC.

TABLE V
KEGG PATHWAYS OF A2-B1-C2-D1 (MOD1)

ID	Term	P Value	Count
KEGG:04080	Neuroactive ligand-receptor interaction	4.19E-16	18
KEGG:04672	Intestinal immune network for IgA production	1.35E-03	3
KEGG:04060	Cytokine-cytokine receptor interaction	2.62E-09	12
KEGG:04062	Chemokine signaling pathway	1.63E-11	12
KEGG:04923	Regulation of lipolysis in adipocytes	1.89E-03	3
KEGG:04924	Renin secretion	2.46E-04	4
KEGG:04657	IL-17 signaling pathway	7.68E-04	4
KEGG:05132	Salmonella infection	5.71E-04	4
KEGG:05134	Legionellosis	1.01E-04	4
KEGG:05146	Amoebiasis	8.66E-04	4

TABLE VI
KEGG PATHWAYS OF A1-B2-C1-D2 (MOD2)

ID	Term	P Value	Count
KEGG:03030	DNA replication	8.21E-04	3
KEGG:03460	Fanconi anemia pathway	2.67E-03	3
KEGG:04110	Cell cycle	1.40E-17	15
KEGG:04115	p53 signaling pathway	3.14E-05	5
KEGG:04218	Cellular senescence	1.25E-06	8
KEGG:04114	Oocyte meiosis	1.87E-07	8
KEGG:04914	Progesterone-mediated oocyte maturation	3.00E-08	8

IV. DISCUSSIONS

Although many researches have discovered that were several cell signal pathways related with the development of CRC, the molecular mechanisms underlying the development of CRC remain unclear. 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 related to CRC were identified based on four stages of DEG sets through pathway enrichment analysis and PPI network modules analysis.

A total number of 35 pathways were found at four stages of CRC. There were several interactive pathways including Pentose and glucuronate interconversions, Ascorbate and aldarate metabolism, Steroid hormone biosynthesis, Retinol metabolism, Drug metabolism, Chemical carcinogenesis, Metabolism of xenobiotics by cytochrome P450, and Fatty acid degradation. Several studies have shown that dietary fats may potentially lead to the development of cancer by regulating the expression of genes in cells to alter metabolism, proliferation and differentiation of cells [23]. Therefore, fatty acid metabolism plays a promoting role in cancer development [24]. Numerous highly polymorphic genes in Cytochrome P450 may affect metabolic ability, which may lead to a different susceptibility to CRC [25]. Some researches shows that Cytochrome P450 was significantly related to CRC development [26]. We considered that closely related metabolic pathways in the network may play an important role in the biological processes of CRC, which contributed to further research on CRC related pathways.

KEGG pathway enrichment analysis of Mod1 mainly enriched Neuroactive ligand-receptor interaction, Intestinal immune network for IgA production, Cytokine-cytokine receptor

interaction, Renin secretion, Chemokine signaling pathway and so on. Chronic intestinal inflammation is one of the important risk factors of CRC. Tumors may generate once the regulation of IgA-secreting cells is imbalanced [27]. Lack of IgA in tumor sites may cause increased inflammation and promote tumor progression, because IgA may decrease bacteria-induced inflammation in tumor sites [28], [29]. In addition, abnormal expression of some cytokines may be related to CRC [30]. For instance, iL-6 is heavily up-regulated in CRC, and it affects epithelial cells in many ways [31]. Moreover, Chemokine signaling affects CRC invasion and/or metastasis by altering tumor microenvironment [32]. Above all, we guess that other enrichment pathways are also associated with the development of CRC.

KEGG pathway enrichment analysis of Mod2 including DNA replication, Fanconi anemia pathway, Cell Cycle, p53 signaling pathway, Cellular senescence, Oocyte meiosis, and Progesterone-mediated oocyte maturation. There is abundant evidence that cell cycle disruption is associated with the development of various cancers. Numerous investigations have indicated that p53 is a tumor suppressor gene [33] and plays a critical role in development of tumors. It contributes to not only control cell cycle [33], but also keep chromosome stability [34] and mitochondrial genetic stability [35], [36]. We consider that other enrichment pathways of Mod2 we obtained are also associated with the development of CRC. More details and evidence are investigated in further studies.

Overall, KEGG pathways we obtained may be associated with pathogenesis of CRC development in this study. Some have been proven to be involved in the development of CRC, and we predict that others might also be related to the development of CRC, which indicates that the method of obtaining related pathways of CRC is useful in this paper. Besides, this method also can be performed to identify pathways in other cancer, such as liver cancer, lung cancer, which also contribute to analyze the dynamic evolution process of cancer.

ACKNOWLEDGMENT

This work was supported by the National Natural Science Foundation of China under [Grant No. 61602386 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.

REFERENCES

[1] Rebecca L Siegel, Kimberly D Miller, and Ahmedin Jemal. Cancer statistics, 2019. *CA: a cancer journal for clinicians*, 69(1):7–34, 2019.

[2] Alexandra Klaus and Walter Birchmeier. Wnt signalling and its impact on development and cancer. *Nature Reviews Cancer*, 8(5):387, 2008.

[3] Ah-Ram Shim, Guang-zhi Dong, Hwa Jin Lee, and Jae-Ha Ryu. Atractylochromene is a repressor of wnt/β-catenin signaling in colon cancer cells. *Biomolecules & therapeutics*, 23(1):26, 2015.

[4] Sangtaek Oh, Jungsug Gwak, Seoyoung Park, and Chung S Yang. Green tea polyphenol egcg suppresses w nt/β-catenin signaling by promoting gsk-3β-and pp2a-independent β-catenin phosphorylation/degradation. *Biofactors*, 40(6):586–595, 2014.

[5] Sihong You, Jiannong Zhou, Senqing Chen, Ping Zhou, Jinghuan Lv, Xiao Han, and Yujie Sun. Ptch1, a receptor of hedgehog signaling pathway, is correlated with metastatic potential of colorectal cancer. *Upsala journal of medical sciences*, 115(3):169–175, 2010.

[6] Bin Xiong, Ling-Ling Gong, Feng Zhang, Ming-Bo Hu, and Hong-Yin Yuan. Tgf-β1 expression and angiogenesis in colorectal cancer tissue. *World journal of gastroenterology*, 8(3):496, 2002.

[7] Piotr G Rychahou, Carlos A Murillo, and B Mark Evers. Targeted rna interference of pi3k pathway components sensitizes colon cancer cells to tnf-related apoptosis-inducing ligand (trail). *Surgery*, 138(2):391–397, 2005.

[8] Christian Gulmann, Katherine M Sheehan, Ronán M Conroy, Julia D Wulfkuhle, Virginia Espina, Michelle J Mullarkey, Elaine W Kay, Lance A Liotta, and Emanuel F Petricoin Iii. 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(4):514–519, 2009.

[9] Xiao-Lan Li, Jianbiao Zhou, Zhi-Rong Chen, and Wee-Joo Chng. P53 mutations in colorectal cancer-molecular pathogenesis and pharmacological reactivation. *World journal of gastroenterology: WJG*, 21(1):84, 2015.

[10] Antonio Russo, Viviana Bazan, Barry Iacopetta, David Kerr, Thierry Soussi, and Nicola Gebbia. 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(30):7518–7528, 2005.

[11] Raymond Cheong, Alex Rhee, Chiaochun Joanne Wang, Ilya Nemenman, and Andre Levchenko. Information transduction capacity of noisy biochemical signaling networks. *science*, 334(6054):354–358, 2011.

[12] Rafael A Irizarry, Bridget Hobbs, Francois Collin, Yasmin D Beazer-Barclay, Kristen J Antonellis, Uwe Scherf, and Terence P Speed. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics*, 4(2):249–264, 2003.

[13] Laurent Gautier, Leslie Cope, Benjamin M Bolstad, and Rafael A Irizarry. affyłanalysis of affymetrix genechip data at the probe level. *Bioinformatics*, 20(3):307–315, 2004.

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

[15] Matthew E Ritchie, Belinda Phipson, Di Wu, Yifang Hu, Charity W Law, Wei Shi, and Gordon K Smyth. limma powers differential expression analyses for rna-sequencing and microarray studies. *Nucleic acids research*, 43(7):e47–e47, 2015.

[16] Gordon K Smyth, Matthew Ritchie, Natalie Thorne, and James Wettenhall. Limma: linear models for microarray data. in bioinformatics and computational biology solutions using r and bioconductor. statistics for biology and health. 2005.

[17] José A Ferreira. The benjamini-hochberg method in the case of discrete test statistics. *The international journal of biostatistics*, 3(1), 2007.

[18] Gabriela Bindea, Bernhard Mlecnik, Hubert Hackl, Pornpimol Charoentong, Marie Tosolini, Amos Kirilovsky, Wolf-Herman Fridman, Franck Pagès, Zlatko Trajanoski, and Jérôme Galon. Cluego: a cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics*, 25(8):1091–1093, 2009.

[19] Gabriela Bindea, Jérôme Galon, and Bernhard Mlecnik. Cluepedia cytoscape plugin: pathway insights using integrated experimental and in silico data. *Bioinformatics*, 29(5):661–663, 2013.

[20] Paul Shannon, Andrew Markiel, Owen Ozier, Nitin S Baliga, Jonathan T Wang, Daniel Ramage, Nada Amin, Benno Schwikowski, and Trey Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*, 13(11):2498–2504, 2003.

[21] Damian Szklarczyk, Andrea Franceschini, Stefan Wyder, Kristoffer Forslund, Davide Heller, Jaime Huerta-Cepas, Milan Simonovic, Alexander Roth, Alberto Santos, Kalliopi P Tsafou, et al. String v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic acids research*, 43(D1):D447–D452, 2014.

[22] Tamás Nepusz, Haiyuan Yu, and Alberto Paccanaro. Detecting overlapping protein complexes in protein-protein interaction networks. *Nature methods*, 9(5):471, 2012.

[23] Donald B Jump and Steven D Clarke. Regulation of gene expression by dietary fat. *Annual review of nutrition*, 19(1):63–90, 1999.

[24] Birgit Hoeft, Jakob Linseisen, Lars Beckmann, Karin Müller-Decker, Federico Canzian, Anika Hüsing, Rudolf Kaaks, Ulla Vogel, Marianne U

Jakobsen, Kim Overvad, et al. Polymorphisms in fatty acid metabolism-related genes are associated with colorectal cancer risk. *Carcinogenesis*, 31(3):466–472, 2009.

[25] Sébastien Küry, Bruno Buecher, Sébastien Robiou-du Pont, Catherine Scoul, Véronique Sébille, Hélène Colman, Claire Le Houérou, Tanguy Le Neel, Jérémie Bourdon, Roger Faroux, et al. Combinations of cytochrome p450 gene polymorphisms enhancing the risk for sporadic colorectal cancer related to red meat consumption. *Cancer Epidemiology and Prevention Biomarkers*, 16(7):1460–1467, 2007.

[26] Gül Özhan, Mine Mutur, Gulcin Ercan, and Buket Alpertunga. Genetic variations in the xenobiotic-metabolizing enzymes cyp1a1, cyp1a2, cyp2c9, cyp2c19 and susceptibility to colorectal cancer among turkish people. *Genetic testing and molecular biomarkers*, 18(4):223–228, 2014.

[27] Xue Wang, Gui-liang Hao, Bo-ya Wang, Chen-chen Gao, Yue-xiu Wang, Li-sheng Li, and Jing-dong Xu. Function and dysfunction of plasma cells in intestine. *Cell & bioscience*, 9(1):26, 2019.

[28] Sabine Le Gouvello, Sylvie Bastuji-Garin, Nijez Aloulou, Hicham Mansour, Marie-Thérèse Chaumette, François Berrehar, Amal Seikour, Antoine Charachon, Mehdi Karoui, Karen Leroy, et al. High prevalence of foxp3 and il17 in mmr-proficient colorectal carcinomas. *Gut*, 57(6):772–779, 2008.

[29] Marie Tosolini, Amos Kirilovsky, Bernhard Mlecnik, Tessa Fredriksen, Stéphanie Mauger, Gabriela Bindea, Anne Berger, Patrick Bruneval, Wolf-Herman Fridman, Franck Pagès, et al. Clinical impact of different classes of infiltrating t cytotoxic and helper cells (th1, th2, treg, th17) in patients with colorectal cancer. *Cancer research*, 71(4):1263–1271, 2011.

[30] LB Trost and JK McDonnell. Important cutaneous manifestations of inflammatory bowel disease. *Postgraduate medical journal*, 81(959):580–585, 2005.

[31] Yuan-Chang Chung and Ya-Fen Chang. Serum interleukin-6 levels reflect the disease status of colorectal cancer. *Journal of surgical oncology*, 83(4):222–226, 2003.

[32] Yoshiro Itatani, Kenji Kawada, Susumu Inamoto, Takamasa Yamamoto, Ryotaro Ogawa, Makoto Taketo, and Yoshiharu Sakai. The role of chemokines in promoting colorectal cancer invasion/metastasis. *International journal of molecular sciences*, 17(5):643, 2016.

[33] Kazufumi Suzuki and Hisahiro Matsubara. Recent advances in p53 research and cancer treatment. *BioMed Research International*, 2011, 2011.

[34] Geetha Achanta, Ryohei Sasaki, Li Feng, Jennifer S Carew, Weiqin Lu, Helene Pelicano, Michael J Keating, and Peng Huang. Novel role of p53 in maintaining mitochondrial genetic stability through interaction with dna pol γ. *The EMBO journal*, 24(19):3482–3492, 2005.

[35] Geng Liu, John M Parant, Gene Lang, Patty Chau, Arturo Chavez-Reyes, Adel K El-Naggar, Asha Multani, Sandy Chang, and Guillermina Lozano. Chromosome stability, in the absence of apoptosis, is critical for suppression of tumorigenesis in trp53 mutant mice. *Nature genetics*, 36(1):63, 2004.

[36] Pablo Iglesias, Antonio Salas, and Jose A Costoya. The maintenance of mitochondrial genetic stability is crucial during the oncogenic process. *Communicative & integrative biology*, 5(1):34–38, 2012.