Identification of key candidate genes for colorectal cancer by bioinformatics analysis

ZHIHUA CHEN1*, YILIN LIN1*, JI GAO2*, SUYONG LIN1, YAN ZHENG1, YISU LIU1 and SHAO QIN CHEN1

1Department of Gastrointestinal Surgery, The First Affiliated Hospital of Fujian Medical University; 2School of Nursing, Fujian Medical University, Fuzhou, Fujian 350004, P.R. China

Received December 27, 2018; Accepted August 16, 2019

DOI: 10.3892/ol.2019.10996

Abstract. Colorectal cancer (CRC) is one of the most common cancers of the digestive tract. Although numerous studies have been conducted to elucidate the cause of CRC, the exact mechanism of CRC development remains to be determined. To identify candidate genes that may be involved in CRC development and progression, the microarray datasets GSE41657, GSE77953 and GSE113513 were downloaded from the Gene Expression Omnibus database. Gene Ontology and Kyoto Encyclopedia of Genes and Genomes were used for functional enrichment analysis of differentially expressed genes (DEGs). A protein–protein interaction network was constructed, and the hub genes were subjected to module analysis and identification using Search Tool for the Retrieval of Interacting Genes/Proteins and Cytoscape. A total of 142 DEGs were identified, with enriched functions and pathways in the ‘cell cycle’, ‘cell proliferation’, ‘the mitotic cell cycle’ and ‘one‑carbon metabolic process’. In addition, 10 hub genes were identified, and functional analysis indicated that these genes are mainly enriched in ‘cell division’, ‘cell cycle’ and functions associated with nucleotide binding processes. Survival analysis demonstrated that DNA topoisomerase II α, cyclin‑dependent kinase 1 and CDC28 protein kinase regulatory subunit 2 may be involved in cancer invasion or recurrence. The DEGs identified in the present study may help explain the molecular mechanisms of CRC development and progression.

Introduction

Colorectal cancer (CRC) ranks third (13.5%) and second (9.5%) among the incidence of malignancies worldwide in male and female patients, respectively, and is a serious hazard to human health (1). Previous studies have demonstrated that the molecular pathogenesis of CRC is mostly caused by genetic mutations (2,3). Numerous studies over the past two decades have reported that genetic mutations are associated with the prognosis and treatment of CRC, and targeted therapies have been developed (4‑7). The progression of CRC is usually accompanied by the activation of the KRAS and BRAF genes and the inhibition of the p53 tumour suppressor gene expression; mutations in these genes are associated with changes in the number and structure of chromosomes (8‑10). However, >15% of sporadic CRCs occur through completely different molecular pathogenesis. For example, serrated precancerous lesions usually manifest as a result of the methylation of the CpG locus and mutation of the gene (11).

The prognosis of CRC is poor due to a lack of effective diagnostic methods at an early stage. Therefore, an effective solution can only be provided for subsequent diagnosis and treatment by better understanding the gene expression of CRC during its occurrence and development and identifying the genes that may be involved in the occurrence and progression of cancer.

With the rapid development of science and technology, microarray technology has become increasingly accurate and has been widely used to explore changes in animal and plant gene expression (12‑14). Microarray technology aids in the discovery of changes in gene expression during cancer development and progression (15‑17). However, it can be difficult to obtain reliable results with single microarray analysis. The present study aimed to identify genetic changes in CRC in three mRNA microarray datasets from the Gene Expression Omnibus (GEO) database by obtaining differentially expressed genes (DEGs) between CRC and normal tissues. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) were used for functional enrichment analysis, and a protein-protein interaction (PPI) network was used to analyse the associations between the DEGs. A total of 142 DEGs and 10 hub genes were identified, which may be candidate biomarkers of CRC.
Materials and methods

Microarray data. GEO (http://www.ncbi.nlm.nih.gov/geo) (18) is a public functional genomic database. Three gene expression datasets (GSE41657, GSE77953, and GSE113513; analysed on Agilent-014850 Whole Human Genome Microarray 4x44K G4112F, Affymetrix Human Genome U133A Array, and Affymetrix Human Gene Expression Array platforms, respectively) were downloaded from GEO. The probes were converted to the corresponding gene symbols based on the annotation information on the platform. The GSE41657 dataset contained 25 CRC and 12 adjacent normal tissue samples. The GSE77953 dataset contained 28 CRC and 13 non-tumour samples. The GSE113513 dataset contained 14 CRC and 14 non-tumour samples.

Identification of DEGs. DEGs between CRC and non-tumour samples were screened using GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r). GEO2R is an interactive web tool that compares two or more GEO datasets to identify DEGs across experimental conditions. Adjusted P-values (adj. P) and Benjamini and Hochberg false discovery rates were applied to balance between statistically significant findings and false-positive limits. Probe sets without corresponding gene symbols were removed, and genes with more than one probe set were averaged. Log [fold change(FC)]>1 and adj. P-value <0.01 were considered to indicate a statistically significant difference.

KEGG and GO enrichment analysis of DEGs. The Database for Annotation, Visualization and Integrated Discovery (DAVID; version 6.7; http://david.ncifcrf.gov) (19) is an online bioinformatics database that integrates biological data and analysis tools and provides gene annotation information and protein data. KEGG is a database resource for understanding advanced functions and biological systems from large-scale molecular data generated by high-throughput experimental techniques (20). GO is a major bioinformatics tool for annotating genes and analysing their biological processes (BPs), molecular functions (MFs) and cellular components (CCs) (21).

To analyse the function of the DEGs, analysis was performed using DAVID. P<0.05 was considered to indicate a statistically significant difference. The results based on the top ten BPs, MFs, CCs and KEGG were visualized.

PPI network construction and module analysis. A PPI network was predicted using the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING; version 10.0; http://string-db.org) database (22). The PPI network of the DEGs was constructed, and interactions with a combined score >0.4 were considered statistically significant. Cytoscape (version 3.4.0) is an open source bioinformatics platform for visualizing molecular interaction networks (23). Cytoscape plug-in Molecular Complex Detection (MCODE; version 1.4.2) provides topology-based clustering for a network (24). Cytoscape was used to draw a PPI network, and MCODE was used to identify the most important modules in the network. The selection criteria were as follows: MCODE score ≥5, degree cut-off =2, node score cut-off =0.2, maximum depth =100, and k score =2. Subsequently, KEGG and GO analysis of the genes in this module was performed using DAVID.

Hub genes selection and analysis. The selection criterion for the hub genes was degree of connectivity ≥10. The functions of the genes were identified using the GeneCards online analysis tool (https://www.genecards.org). The BP analysis and visualization of the hub genes were performed using the cBioPortal (http://www.cbioportal.org) online platform (25). The functions of the hub genes were identified using GO and KEGG analysis. Hierarchical clustering of the hub genes was constructed using the University of California Santa Cruz Cancer Genomics Browser (http://genome-cancer.ucsc.edu) (26). The overall survival analysis was performed using Kaplan-Meier curves in cBioPortal. The log-rank test was used to for statistical analysis. The expression profiles of DNA topoisomerase I (TOP2A) and phosphoribosylaminomimidazole carboxylase (PAICS) were analysed using the Oncomine database (http://www.oncomine.com). The expression levels of genes in normal tissues and CRC tissues were also analysed using the Oncomine database (27).

Results

Determination of DEGs in CRC. DEGs were identified by standardized microarray results (GSE41657, GSE77953 and GSE113513). A total of 142 genes overlapped in the three datasets, as demonstrated in the Venn diagram (Fig. 1A).

PPI network and module analysis. A PPI network of the 142 identified DEGs was constructed (Fig. 1B). A significantly enriched module was obtained using Cytoscape software (Fig. 1C).

GO and KEGG enrichment analysis of the DEGs. DAVID was used for the functional enrichment analysis of DEGs. The results indicated that the DEGs were mainly enriched in ‘cell proliferation’, ‘G2/M transition of mitotic cell cycle’ and ‘one-carbon metabolism’ BPs (Fig. 2A); ‘cytoplasm’, ‘cytosol’ and ‘extracellular exosome’ cellular components (CCs) (Fig. 2B); and ‘protein binding’ and ‘ATP binding’ molecular functions (MFs) (Fig. 2C). KEGG pathway analysis revealed strong enrichment in ‘biosynthesis of antibiotics’, ‘purine metabolism’, ‘pancreatic secretion’ and ‘mineral absorption’ (Fig. 2D).

The functional results of the genes in the most significant module indicated that these genes were mainly enriched in processes associated with cell cycle progression, ATP binding, nucleotide binding, the p53 signalling pathway and oocyte meiosis (Table I).

Hub genes screening and analysis. Genes with a degree of connectivity ≥10 were identified as the hub genes. The names, abbreviations and functions of these genes are presented in Table II. The cBioPortal online platform was used to construct a hub gene network (Fig. 3A), and functional enrichment analysis of DEGs was performed using DAVID (Table III). The heat map constructed using the UCSC Cancer Genomics Browser indicated that the hub genes may be used to distinguish CRC tissue samples from normal intestinal cancer tissue samples (Fig. 3B). Overall and progression-free survival
based on alterations in the hub genes was subsequently analysed using Kaplan-Meier curves. Alterations in TOP2A, cyclin-dependent kinase 1 (CDK1) and CDC28 protein kinase regulatory subunit 2 (CKS2) in patients with CRC were associated with a poor overall survival rate (Fig. 4A). However, patients with alterations in TOP2A and CKS2 did not exhibit significant differences in progression-free survival (Fig. 4B).

Among these genes, TOP2A exhibited a node degree of 34, and PAICS exhibited the highest degree of connectivity among the hub genes (34 and 26, respectively). These two genes may serve important roles in the occurrence or development of CRC. Based on the survival analysis, alterations in PAICS exhibited no statistically significant differences in overall and progression-free survival in patients with CRC (overall survival P=0.986; progression-free survival P=0.918; Fig. 4). Overall survival rates were significantly different for patients with TOP2A alterations, but progression-free survival was not significantly different (overall survival P=0.002; progression-free survival P=0.703; Fig. 4). The expression profiles of TOP2A and PAICS in human tissues were

Figure 1. DEGs between CRC and non-tumour tissues. (A) Venn diagram of the DEGs in the three data sets (GSE41657, GSE77953, and GSE113513). Log fold change >1 and adjusted P-value <0.01 were the conditions for screening DEGs. A total of 142 DEGs were screened. (B) PPI network of 142 DEGs. Yellow indicates the densest module. (C) The relationships among the most important module genes in the PPI network. DEGs, differentially expressed genes; CRC, colorectal cancer; PPI, protein-protein interaction.
analysed using the Oncomine database; TOP2A and PAICS mRNA was upregulated in various cancer tissues compared with normal tissues (Fig. 5A and B). Further analysis revealed significant increases in the expression of TOP2A and PAICS in CRC (Fig. 5C and D).

Discussion

CRC is a common malignant tumour of the digestive tract, and its mortality rate (9.2%) ranks second among all types of cancer (1). The main causes of CRC include dietary and environmental factors, as well as genetic mutations (3). The simultaneous methylation of the CpG site and mutation of the BRAF gene are also important factors in the development and progression of CRC (11). Failure to detect CRC early may be one of the reasons for poor prognosis in patients. Therefore, there is an urgent need for efficient diagnostic and therapeutic methods.

In the present study, DEGs between CRC and non-cancerous tissues were obtained from three mRNA microarray datasets.
The functions of the DEGs were identified by GO and KEGG enrichment analysis. The results demonstrated that the DEGs were mainly enriched in the cell cycle, proliferation, mitotic cell cycle and carbon metabolism. Previous studies have reported that dysregulation of the cell cycle and the mitotic cell cycle serves an important role in the development and progression of CRC (28-31). In a study by Chamberlain et al (32), functional B vitamin was used to assess the association of carbon metabolism with CRC, and a similar association between total B vitamin status and CRC risk was identified. Ducker (33) et al also revealed that the single-carbon metabolism of the cytosol can support tumourigenesis. These findings are consistent with the results of the present study. GO cluster and KEGG analysis in the present study also revealed that the changes in the most important module were mainly enriched in processes associated with cell cycle progression, ATP and

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
<th>No. of genes</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0007049</td>
<td>Cell cycle</td>
<td>11</td>
<td>8.42x10^-10</td>
</tr>
<tr>
<td>GO:0000279</td>
<td>M phase</td>
<td>9</td>
<td>6.24x10^-10</td>
</tr>
<tr>
<td>GO:0022403</td>
<td>Cell cycle phase</td>
<td>9</td>
<td>3.84x10^-9</td>
</tr>
<tr>
<td>GO:0022402</td>
<td>Cell cycle process</td>
<td>9</td>
<td>4.38x10^-8</td>
</tr>
<tr>
<td>GO:0043228</td>
<td>Non-membrane-bounded organelle</td>
<td>11</td>
<td>1.83x10^-5</td>
</tr>
<tr>
<td>GO:0015630</td>
<td>Microtubule cytoskeleton</td>
<td>9</td>
<td>1.17x10^-4</td>
</tr>
<tr>
<td>GO:0044430</td>
<td>Cytoskeletal part</td>
<td>9</td>
<td>8.45x10^-7</td>
</tr>
<tr>
<td>GO:0005524</td>
<td>ATP binding</td>
<td>6</td>
<td>2.91x10^-3</td>
</tr>
<tr>
<td>GO:0032559</td>
<td>Adenyl ribonucleotide binding</td>
<td>6</td>
<td>3.10x10^-3</td>
</tr>
<tr>
<td>GO:0030554</td>
<td>Adenyl nucleotide binding</td>
<td>6</td>
<td>3.91x10^-3</td>
</tr>
<tr>
<td>hsa04115</td>
<td>p53 signalling pathway</td>
<td>3</td>
<td>1.72x10^-3</td>
</tr>
<tr>
<td>hsa04114</td>
<td>Oocyte meiosis</td>
<td>3</td>
<td>4.44x10^-3</td>
</tr>
<tr>
<td>hsa04110</td>
<td>Cell cycle</td>
<td>3</td>
<td>5.71x10^-3</td>
</tr>
</tbody>
</table>

GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

Table II. Functional roles of the 10 hub genes with degree of connectivity ≥10.

<table>
<thead>
<tr>
<th>No.</th>
<th>Gene symbol</th>
<th>Full name</th>
<th>Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>TOP2A</td>
<td>DNA Topoisomerase II α</td>
<td>Target for anticancer agents; mutations are associated with drug resistance</td>
</tr>
<tr>
<td>2</td>
<td>PAICS</td>
<td>Phosphoribosylaminomimidazole carboxylase</td>
<td>Involved in de novo synthesis of purine nucleotides</td>
</tr>
<tr>
<td>3</td>
<td>CDK1</td>
<td>Cyclin-dependent kinase 1</td>
<td>Regulates cell cycle progression, apoptosis and carcinogenesis of tumour cells</td>
</tr>
<tr>
<td>4</td>
<td>CKS2</td>
<td>CDC28 protein kinase regulatory subunit 2</td>
<td>Binds to the catalytic subunit of cyclin-dependent kinases; essential for their biological function</td>
</tr>
<tr>
<td>5</td>
<td>CKAP2</td>
<td>Cytoskeleton associated protein 2</td>
<td>Possesses microtubule stabilizing properties; involved in regulating aneuploidy, cell cycle and cell death in a p53/TP53-dependent manner</td>
</tr>
<tr>
<td>6</td>
<td>CEP55</td>
<td>Centrosomal protein 55</td>
<td>Promotes the proliferation of lung, breast and thyroid cancer</td>
</tr>
<tr>
<td>7</td>
<td>VEGFA</td>
<td>Vascular endothelial growth factor A</td>
<td>Induces endothelial cell proliferation, promotes cell migration, inhibits apoptosis and induces blood vessel permeabilization</td>
</tr>
<tr>
<td>8</td>
<td>PHLPP2</td>
<td>PH domain and leucine rich repeat protein phosphatase 2</td>
<td>Regulates Akt and PKC signalling.</td>
</tr>
<tr>
<td>9</td>
<td>RRM2</td>
<td>Ribonucleotide reductase M2 polypeptide</td>
<td>Catalyses the biosynthesis of deoxyribonucleotides; inhibits Wnt signalling</td>
</tr>
<tr>
<td>10</td>
<td>NEK2</td>
<td>NIMA (never in mitosis gene a)-related kinase 2</td>
<td>Involved in the regulation of mitosis</td>
</tr>
</tbody>
</table>

The functions of the DEGs were identified by GO and KEGG enrichment analysis. The results demonstrated that the DEGs were mainly enriched in the cell cycle, proliferation, mitotic cell cycle and carbon metabolism. Previous studies have reported that dysregulation of the cell cycle and the mitotic cell cycle serves an important role in the development and progression of CRC (28-31). In a study by Chamberlain et al (32), functional B vitamin was used to assess the association of carbon metabolism with CRC, and a similar association between total B vitamin status and CRC risk was identified. Ducker (33) et al also revealed that the single-carbon metabolism of the cytosol can support tumourigenesis. These findings are consistent with the results of the present study. GO cluster and KEGG analysis in the present study also revealed that the changes in the most important module were mainly enriched in processes associated with cell cycle progression, ATP and
A total of 10 DEGs were selected as the hub genes. Among them, TOP2A and PAICS were the top two nodes with the highest degree of connectivity. TOP2A is involved in DNA replication, transcription and chromosome segregation, and is essential for tumorigenesis and cancer development (34). TOP2A has been demonstrated to be associated with chemoresistance and tumour recurrence and has been recognized as a target for anticancer drugs (35,36). In addition, TOP2A is highly expressed in lung, colon and breast cancer, involved in the inhibition of apoptosis, proliferation and chemoresistance of CRC and may be considered an important biomarker for the diagnosis, prognosis and treatment of tumours (37–39).

In the present study, the PPI network indicated that TOP2A interacted directly with CDK1, ribonucleoside-diphosphate reductase subunit M2 (RRM2) and CKS2, indicating a key role for TOP2A in CRC. PAICS is involved in purine nucleotide synthesis; a previous study has demonstrated that high expression of PAICS in lung and prostate cancer is associated with an altered metabolic state of cells, apoptosis inhibition in cancer cells and enhanced cancer cell invasion (40). In addition, PAICS has been reported to promote tumorigenesis and progression in breast and bladder cancer (41,42). The results of the analysis in the Oncomine database in the present study demonstrated that TOP2A and PAICS expression was significantly upregulated in CRC compared with that in normal tissues; however, to the best of our knowledge, there are currently no studies focusing on the association between PAICS expression and CRC. The association between TOP2A and PAICS and overall and disease-free survival was also analysed; changes in TOP2A were significantly associated with overall survival but were independent of disease-free survival. Changes in PAICS exhibited a reduction in overall and disease-free survival; however, these observations were

<table>
<thead>
<tr>
<th>Term</th>
<th>Description</th>
<th>No. of genes</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0007049</td>
<td>Cell cycle</td>
<td>5</td>
<td>6.24x10⁻⁴</td>
</tr>
<tr>
<td>GO:0051301</td>
<td>Cell division</td>
<td>4</td>
<td>5.30x10⁻⁴</td>
</tr>
<tr>
<td>GO:0000279</td>
<td>M phase</td>
<td>4</td>
<td>7.29x10⁻⁴</td>
</tr>
<tr>
<td>GO:0015630</td>
<td>Microtubule cytoskeleton</td>
<td>5</td>
<td>1.06x10⁻⁴</td>
</tr>
<tr>
<td>GO:0044430</td>
<td>Cytoskeletal part</td>
<td>5</td>
<td>8.92x10⁻⁴</td>
</tr>
<tr>
<td>GO:0005856</td>
<td>Cytoskeleton</td>
<td>5</td>
<td>3.63x10⁻³</td>
</tr>
<tr>
<td>GO:0005524</td>
<td>ATP binding</td>
<td>4</td>
<td>3.62x10⁻²</td>
</tr>
<tr>
<td>GO:0032559</td>
<td>Adeny ribonucleotide binding</td>
<td>4</td>
<td>3.75x10⁻²</td>
</tr>
<tr>
<td>GO:0030554</td>
<td>Adenyl nucleotide binding</td>
<td>4</td>
<td>4.29x10⁻²</td>
</tr>
<tr>
<td>hsa04115</td>
<td>p53 signalling pathway</td>
<td>2</td>
<td>3.96x10⁻²</td>
</tr>
</tbody>
</table>

**Table III. GO and KEGG pathway enrichment analysis of hub genes.**

GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes.

![Figure 3. Interaction network of hub genes and hierarchical cluster analysis. (A) Analysis of the hub genes and their co-expressed genes using cBioPortal. (B) Hierarchical clustering of the hub genes. In the sample type column, brown samples are non-cancerous tissue samples, and blue samples are CRC tissue samples. In the gene expression columns, red indicates upregulated expression and blue indicates downregulated expression. CRC, colorectal cancer.](image)
not statistically significant. These results may require further research for verification.

CDK1 is a non-redundant cyclin-dependent kinase that serves an important role in mitosis (43,44). Perturbations in chromosomal stability and aspects of S phase and G2/M control mediated by CDK2 and CDK1 are pivotal tumorigenic events (45). A previous study has demonstrated that CDK1 is required for the survival of cells overexpressing MYC, and CDK1 has a therapeutic effect in the treatment of human malignancies that overexpress MYC (46). The combination of CKS2 and CDK is essential for promoting cancer cell metastasis in diseases such as colon cancer (47,48). Vascular endothelial growth factor A (VEGFA) is a class of cytokines called antigen growth factors that stimulate the formation of new blood vessels; tumour angiogenesis is primarily dependent on VEGFA-driven responses (49). Anti-VEGF-A treatment can reduce Treg proliferation in CRC patients, and VEGFA inhibitors have been used to treat CRC (50,51). In the present study revealed, the expression levels of CDK11, CSK2 and VEGFA in CRC were analysed; the results revealed high expression of these genes in CRC, which was consistent with previous studies. In addition, high expression levels of CDK1 and CKS2 were associated with poor survival.

CKAP2 is a spindle-associated protein that is degraded during mitosis (52,53). Upregulation of CEP55 promotes the metastasis of several cancers, such as lung adenocarcinoma, breast and anaplastic thyroid cancer (54-56). RRM2, also known as the ribonucleotide reductase small subunit, promotes cyclin F degradation to maintain a balanced dNTP pool and genomic stability (57). PHLPP2 is a protein phosphatase involved in the regulation of Akt and PKC signalling (58,59). Previous studies have reported that microRNA-224 promotes the proliferation and tumour growth of human CRC cells by inhibiting PHLPP1 and PHLPP2 (60) and acts as a tumour suppressor for lung and breast cancer (61,62). The results of the present study also demonstrated that it is downregulated in CRC tissues. NEK2 is a serine/threonine kinase that is a key regulator of mitosis in cells (63). Upregulation of NEK2 is associated with the development of cancer, such as breast (64), ovarian (65) and colorectal (66,67) cancer. Overexpression of NEK2 promotes the invasion and metastasis of tumour cells and is considered a potential biomarker in nasopharyngeal carcinoma and liver cancer (68,69). The results of the present study indicated that the aforementioned hub genes may be distinguishing factors between CRC and non-tumour tissue samples and candidate tumour biomarkers. In addition, alterations in TOP2A, CDK1 and CKS2 are associated with poor survival, suggesting that these genes may serve an important role in the development, progression or recurrence of cancer.

Figure 4. Analysis of overall survival and progression-free survival of four hub genes. (A) Overall survival and (B) progression-free survival analyses based on alterations in the hub genes were performed using the cBioPortal online platform. P<0.05 was considered to indicate a statistically significant difference. TOP2A, DNA topoisomerase IIα; PAICS, phosphoribosylaminomimidazole carboxylase; CDK1, cyclin-dependent kinase 1; CKS2, CDC28 protein kinase regulatory subunit 2.
Figure 5. Expression of TOP2A and PAICS in cancer and normal tissues. (A and B) The Oncomine online analysis platform was used to analyse the expression of (A) TOP2A and (B) PAICS in human cancer and normal tissues. (C) Heat map of TOP2A expression in clinical CRC and normal tissues. 1. Colorectal Carcinoma vs. Normal. Hong Colorectal. 2. Rectosigmoid Adenocarcinoma vs. Normal. Kaiser Colon. 3. Colon Adenocarcinoma vs. Normal. Notterman Colon. 4. Colorectal Carcinoma vs. Normal. Skrzypczak Colorectal. (D) Heat map of PAICS expression in clinical CRC tissues and normal tissues. Caecum Adenocarcinoma, Colon Adenocarcinoma, Colon Mucinous Adenocarcinoma and Rectal Adenocarcinoma vs. Normal. TCGA Colorectal. TOP2A, DNA topoisomerase IIα; PAICS, phosphoribosylaminomimidazole carboxylase; CRC, colorectal cancer.
In summary, the aim of the present study was to identify DEGs that may be involved in the development or progression of CRC. The results demonstrated that TOP2A and CDK1 may be involved in the survival prognosis of patients with CRC. The present study also revealed that these genes may be biomarkers for the diagnosis of CRC. However, the present study had limitations, and other databases such as The Cancer Genome Atlas, as well as in vivo and in vitro experiments may be needed to clarify the biological functions of these genes in CRC.

Acknowledgements
Not applicable.

Funding
The present study was supported by the Fujian Provincial Key Medical Specialist Construction Project (grant no. 2016-SLCZD), the Training Project of Young Talents in the Health System of Fujian Province (grant no. 2016-ZQN-45) and the Fujian Medical University Start up Fund for Scientific Research (grant no. 2017-XQ1061).

Availability of data and materials
The datasets used and analysed during the present study are available from the corresponding author on reasonable request.

Authors' contributions
SQC and ZHC designed the study and obtained funding YLL drafted the manuscript and supervised the study. YLL participates in the design of the study, obtaining data and writing manuscripts. YZ, YSL, JG and SYL collected and analysed the data. All authors read and approved the final manuscript.

Ethics approval and consent to participate
Not applicable.

Patient consent for publication
Not applicable.

Competing interests
The authors declare that they have no competing interests.

References


This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.