European Review for Medical and Pharmacological Sciences 2017; 21: 2114-2129

Comprehensive analysis of microRNA/mRNA signature in colon adenocarcinoma

J.-Y. WANG[1,2], C.-L. WANG[3], X.-M. WANG[2], F.-J. LIU[4]

[1]Department of Colorectal and Anal Surgery, Qilu Hospital of Shandong University, Ji'nan, China
[2]Department of Colorectal and Anal Surgery, Jining No. 1 People's Hospital, Jining, China
[3]Department of Gastrointestinal Surgery, Jining No. 1 People's Hospital, Jining, China
[4]Department of General Surgery, Qilu Hospital of Shandong University, Ji'nan, China

Abstract. – OBJECTIVE: The goal of our study was to identify the regulatory mechanisms of gene expression mediated by miRNAs and DNA methylation in colon adenocarcinoma (COAD).

MATERIALS AND METHODS: The miRNAs and mRNAs expression and DNA methylation data of COAD and adjacent normal tissues were obtained from The Cancer Genome Atlas (TCGA) database. Based on the differentially expressed miRNAs and mRNAs, miRNA-mRNA pairs were obtained by correlation analysis and prediction algorithms. Finally, COAD-specific miRNA-mRNA regulatory network was generated. Additionally, the biological functions of miRNA targets were further revealed by GO and KEGG enrichment analysis. Besides, the correlation analysis between gene expression and DNA methylation was also performed after differential analysis.

RESULTS: We identified 55 differentially expressed miRNAs and 1291 differentially expressed mRNAs in COAD compared with adjacent normal tissues. We observed a global miRNA up-regulation in tumors. A total of 58 miRNA-mRNA pairs were not only predicted by algorithms but also negatively correlated. The increased expression of has-mir-141, -19a, -20a 19b-1, 19b-2, 16, 590 and -335 were closely associated with the carcinogenesis of COAD. Functional enrichment analysis showed that the miRNA targets were significantly enriched in pancreatic secretion, salivary secretion, gastric acid secretion and bile secretion. Regarding the regulatory role of DNA methylation, we identified 11 genes whose expressions were negatively correlated with DNA methylation level. Among those genes, MSX1 and KRT7 were down-regulated and hypermethylated in COAD compared with adjacent normal tissues.

CONCLUSIONS: These eight miRNAs (has-mir-141, -19a, -20a 19b-1, 19b-2, 16, 590 and -335) and two genes (MSX1 and KRT7) may play a role in the process of COAD. These findings highlighted the potential regulatory mechanisms of miRNA and DNA methylation on mRNA expression in COAD carcinogenesis.

Key Words
Colon adenocarcinoma, microRNAs, DNA methylation.

Introduction

In terms of occurrence, Colon cancer ranks third, and it is also the fifth leading cause of cancer-related deaths[1]. As the most frequent subtype of colon cancer, colon adenocarcinoma (COAD) is increasingly occurring in China. COAD develops through a multi-stage process via transforming normal mucosa to adenoma and then to carcinoma[2]. At present, optical colonoscopy is considered as the gold standard for COAD screening. However, optical colonoscopy is costly and invasive with a low rate of compliance and unfavorable prognosis[3,4]. Therefore, it is important to develop the non-invasive diagnostic markers for detection of COAD.

Up to now, potential diagnostic and prognostic biomarkers of colon cancer have been identified. Dysregulation of fibulin-3, HSP70 and Caspase-3 were found in colon cancer, which can serve as diagnostic biomarkers for colon cancer[5-8]. Several prognostic factors of colon cancer including stathmin 1, p21, p53, Cyclooxigenase-2, E-cadherin, CD44 and the Ki-67 protein have also been identified[9-12]. Recently, some long noncoding RNA was used as biomarkers in colon cancer[9,13,14].

The aberrant microRNA (miRNA) expression also has a functional role in the development and progression of COAD by acting as oncogenes and tumour suppressors[15-17]. Moreover, miRNAs can target hundreds of mRNAs and disturb cell signaling pathways. Particularly, the circulating miRNAs have been widely considered as the non-invasive early detection biomarkers[18-21]. Previous microarray studies have shown a bunch of miRNA/mRNA deregulated pairs involved in COAD tumorigenesis, which provided new clues for the pathogenesis of COAD[22]. In addition, accumulating evidence demonstrated that alterations in DNA methylation also play important roles in tumor formation and progression

Corresponding Author: Fengjun Liu, MD; e-mail: fengjunliudoctor@sina.com

by affecting gene expression[23-26]. Understanding epigenetic modification of COAD can facilitate the etiological research of this disease.

The Cancer Genome Atlas (TCGA) Data Portal contains clinical information, genomic characterization data, and high level sequence analysis on cases and controls of COAD. Consequently, TCGA can provide unprecedented opportunities to reveal the molecular mechanisms of COAD. In the present study, we obtained the miRNA, mRNA expression and DNA methylation data of a large number of patients with COAD from the TCGA database. Based on these data, we performed a comprehensive analysis by using a bioinformatics approach to discover the regulatory roles of miRNAs and DNA methylation in the development of COAD.

Materials and Methods

Integrated Profiles in TCGA

TCGA is a central bank for multidimensional experimental cancer data, providing unprecedented opportunities to reveal molecular mechanisms of cancer. At the time of January 5, 2016, TCGA consisted of 459 COAD patients with clinical data. In this work, the mRNA expression, miRNA expression, DNA methylation data and clinical data were downloaded from TCGA (https://tcga-data.nci.nih.gov/tcga/tcgaHome2.jsp). The patients with other malignancy history or neoadjuvant treatment history were excluded. Only patients who were diagnosed as COAD histologically were included.

Identifying miRNAs and mRNAs to Distinguish Tumour from Normal

Based on the read count of each sample, the differentially expressed miRNAs and mRNAs in COAD compared to adjacent tissues were calculated via R-bioconductor package DESeq2[27]. We performed multiple comparisons by using the Benjamini and Hochberg method to obtain the false discovery rate (FDR)[28]. The threshold for the expression of differentially expressed miRNAs was FDR<0.0001 and $|\mathrm{Log_2}\text{fold change}| > 4$. For differentially expressed mRNAs, the threshold was defined as FDR<0.0001 and $|\mathrm{Log_2}\text{fold change}| > 2$. In search of a reasonable amount of differently expressed miRNAs and mRNAs, the absolute Log2 fold change values were determined by the Volcano plots of miRNAs and mRNAs (Supplementary Figure S1 and Figure S2).

Identifying mRNA Targets of miRNA

The pairwise Pearson correlation coefficients between differentially expressed miRNAs and differentially expressed genes were calculated. Statistical significance was defined as $p<0.05$ and $r< -0.3$[29]. Moreover, the putative targets of differentially expressed miRNAs were predicted by six bioinformatic algorithms (RNA22, miRanda, miRDB, miRWalk, PICTAR2 and Targetscan), and the targets recorded by more than 4 algorithms were selected. All the miRNA-target pairs were finally obtained, and the miRNA-target regulatory network was constructed, which was visualized using Cytoscape software (http://www.cytoscape.org/).

Functional Annotation of miRNA Targets

To uncover the biological functions and detect the potential pathways of miRNA targets, the online software GeneCodis[30] was used to perform the functional annotation. The analysis of Gene Ontology (GO) classification (molecular functions, biological processes and cellular component)[31] and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment[32] were performed. Statistical significance was defined as FDR<0.05.

Differential Methylation Analysis

By using COHCAP package in R (https://sourceforge.net/projects/cohcap/), the differentially methylated genes in COAD compared to adjacent normal tissues were identified[33]. Theses differentially methylated genes were involved with differentially methylated sites and regions which can regulate the expression of their downstream genes. FDR<0.05 was indicated the significant difference.

Correlation Analysis Between Differential mRNAs and DNA Methylation

To analyze the association between gene expression and DNA methylation, we identified the genes which were not only differentially expressed, but also differentially methylated. Since DNA methylation negatively influences the corresponding gene expression, the differentially expressed genes whose expression was inversely correlated with DNA methylation level were identified with COHCAP package in R.

Confirmation of Differentially Expressed miRNAs and mRNAs

The tumor and adjacent normal tissues were obtained from three patients who were diagnosed as COAD. The resected tissues were immediately

frozen in liquid nitrogen. We obtained the informed consent and the approval from the Ethics Committee of our hospital.

Total RNA was extracted with the Trizol reagent (Invitrogen, Shanghai, China). 1 μg RNA was used to synthesize cDNA using SuperScript® III Reverse Transcriptase (Invitrogen, Shanghai, China). Quantitative PCR was performed with Power SYBR® Green PCR Master Mix (Applied Biosystems, Carlsbad, CA, USA) in ABI7500 real-time PCR system (Applied Biosystems, Carlsbad, CA, USA). The reverse transcriptions of miRNAs were performed using the miScript II RT Kit (Qiagen, Hilden, Germany). Quantitative PCR was performed with miScript SYBR Green PCR Kit (Qiagen, Hilden, Germany). Relative gene expression was analyzed using $2^{-\Delta\Delta Ct}$ method. The human 18srRNA and U6 were used as endogenous controls for mRNA and miRNA expression in analysis.

In addition, a miRNA expression dataset of COAD (GSE83924) derived from the Gene Expression OMBINUS (GEO) was used to verify the expression of miRNAs as well. To evaluate the efficiency of differentially expressed miRNAs as potential diagnostic biomarkers of COAD, the Receiver Operator Characteristic (ROC) curves between COAD patients and healthy controls were obtained via R-bioconductor package pROC.

Results

Identification of Differentially miRNAs and mRNAs

Clinical data of 459 patients with COAD were obtained from TCGA data portal. Based on clinical data, 62 patients with other malignancy history or neoadjuvant treatment were excluded. Moreover, 60 patients who were diagnosed as colon mucinous adenocarcinoma, not available or discrepancy were also excluded. According to the inclusion criteria, 337 COAD patients were enrolled in this study. The miRNA expression was performed on the IlluminaHiSeq-miRNASeq platform measuring the expression of 1046 miRNAs. According to principal component analysis (PCA, Supplementary Figure S3 and Figure S4), 9 discrete samples were eliminated, leaving a final cohort of 193 tumors and 5 adjacent normal tissues of the same patients. For MiRNAs with read count value=0 in more than 20% tumors (n=40) or in more than 20% adjacent samples (n=1), they were considered to be difficultly detected miRNAs. After filtering the miRNAs difficultly detected, 407 miRNAs were retained.

The mRNA expression was performed on the IlluminaHiseq_RNASeqV2 platform measuring the expression of 20531 genes. The final cohort was consisted of 219 tumors and 29 adjacent tissues of the same patients (Supplementary Figure S5). MRNAs with read count value=0 in more than 20% tumors (n=43) or in more than 20% adjacent samples (n=5) were considered to be difficult to detect. After filtering the mRNAs difficultly detected, 16142 mRNAs were retained. The final cohort contained 219 cases and 29 normal samples.

Differentially Expressed miRNAs and mRNAs in COAD

There were 55 differentially expressed miRNAs (15 down-regulated and 40 up-regulated miRNAs) between COAD and normal tissues with FDR<0.0001 and |Log2 fold change|>4 (Figure 1). Moreover, 1291 differentially expressed mRNAs (803 down-regulated and 488 up-regulated mRNAs) between COAD and normal tissues were identified with FDR<0.0001 and |Log2 fold change|>2. The heat-map of top 100 differentially expressed mRNAs was displayed in Figure 2.

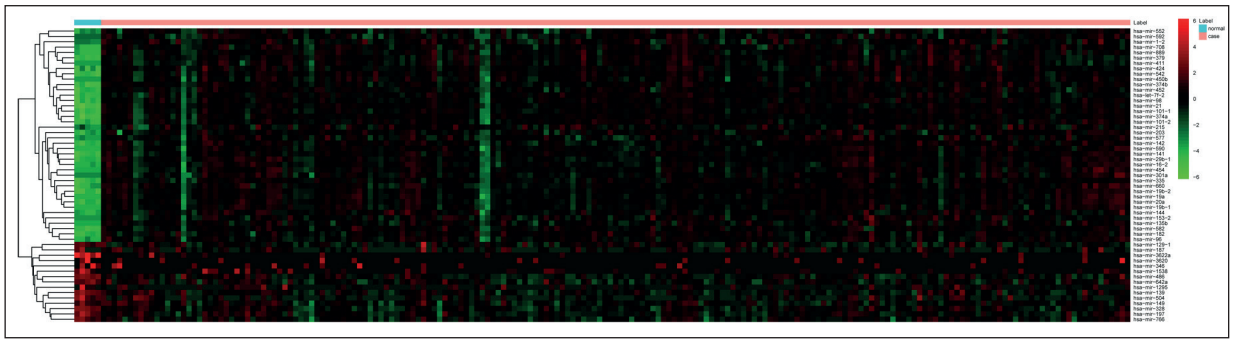

Figure 1. The heatmap of differentially expressed miRNAs in COAD compared to adjacent tissues.

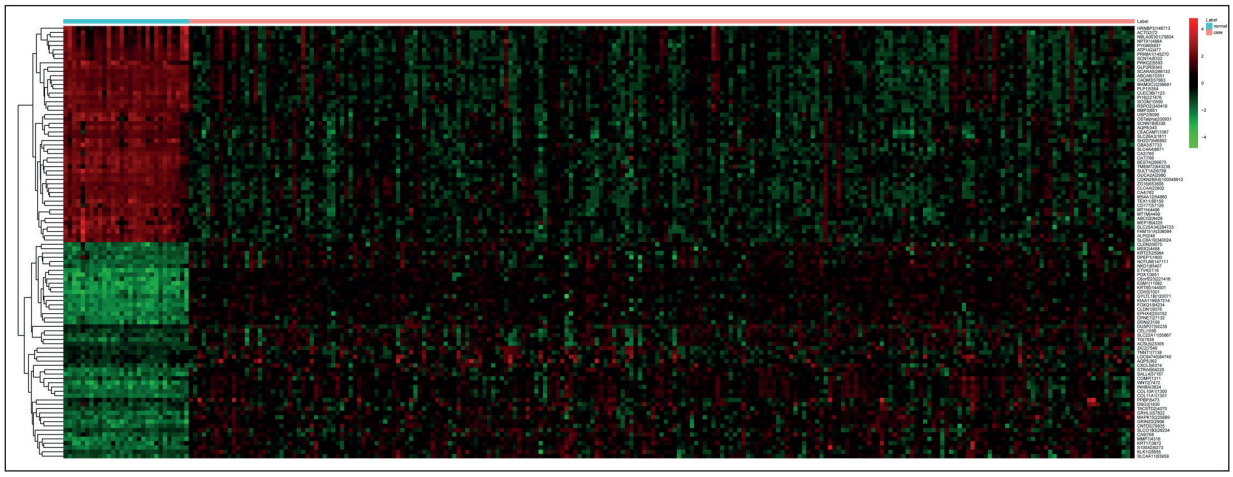

Figure 2. The heat map of top 100 differentially expressed mRNAs in COAD compared to adjacent tissues.

miRNA-mRNA Interactions in COAD

MiRNAs are negative regulators of the expression of targets were negatively associated with miRNAs[34]. According to the miRNA-mRNA expression correlation analysis, we obtained 778 miRNA-mRNA pairs which were negatively correlated (p<0.05, r<-0.3)[29]. Moreover, we obtained 1388 miRNA-mRNA pairs predicted by more than four algorithms, consisting of 472 down-regulated miRNA-mRNA pairs (Supplementary Figure S6) and 1916 up-regulated miRNA-mRNA pairs (Supplementary Figure S7). Taken together, we obtained 2623 miRNA-mRNA pairs in total, in which 58 miRNA-mRNA pairs were not only predicted by algorithms but also negatively correlated (Table I). Based on the miRNA-mRNA pairs, the miRNA-mRNA regulatory network was generated, which consisted of 8 miRNAs (hsa-mir-141, hsa-mir-16-2, hsa-mir-19a, hsa-mir-19b-1, hsa-mir-19b-2, hsa-mir-20a, hsa-mir-335 and hsa-mir-590) and 47 mRNA targets (Figure 3). In the network, the hsa-mir-141 regulated the most target genes.

Functional Enrichment Analysis of miRNA Targets

GO enrichment analysis (Figure 4) revealed that the miRNA targets were significantly enriched in cell adhesion (FDR=2.76E-11), positive regulation of cell proliferation (FDR= 1.61E-07), cell differentiation (FDR=1.28E-06) and negative regulation of cell proliferation (FDR=7.62E-06) (data not shown). KEGG pathway enrichment showed that several pathways associated with digestion were enriched, including pancreatic secretion (FDR= 1.09E-06), salivary secretion (FDR= 5.58E-05), gastric acid secretion (FDR= 6.66E-05) and bile secretion (FDR=0.00149901) (Table II and Figure 5)

Table I. miRNA-mRNA pairs not only predicted by algorithms but also had the negative correlations.

miRNA	Regulation (miRNAs)	Count of targets	Target mRNAs
hsa-mir-141	Up	19	CACNA2D1,PEG3,PGR,SCN2B,SCN4B,CXCL12,SLIT3,TLL1, TNS1,ANGPTL1,DAAM2,PAMR1,NECAB1,LYNX1,KIAA1644, ZDHHC15,ATP6V0D2,RSPO2,KIAA2022
hsa-mir-16-2	Up	14	MEIS1,MYLK,PGR,TCF21,PPAP2A,DCLK1,P2RY14,EDIL3, ARHGAP20,NECAB1,CLIP4,RERG,CNTN4,NEGR1
hsa-mir-19a	Up	3	JAM2,CLIP4,NEGR1
hsa-mir-19b-1	Up	1	HPGDS
hsa-mir-19b-2	Up	5	CFL2,DCLK1,JAM2,CLIP4,NEGR1
hsa-mir-20a	Up	2	FGL2,JAM2
hsa-mir-335	Up	12	BMP2,PTPRH,GCNT3,SMPDL3A,PLAC8,TMCC3,TMCC3, KIAA1211,MMP28,CCDC68,ZBTB7C,EIF4E3
hsa-mir-590	Up	2	SIGLEC6,NEGR1

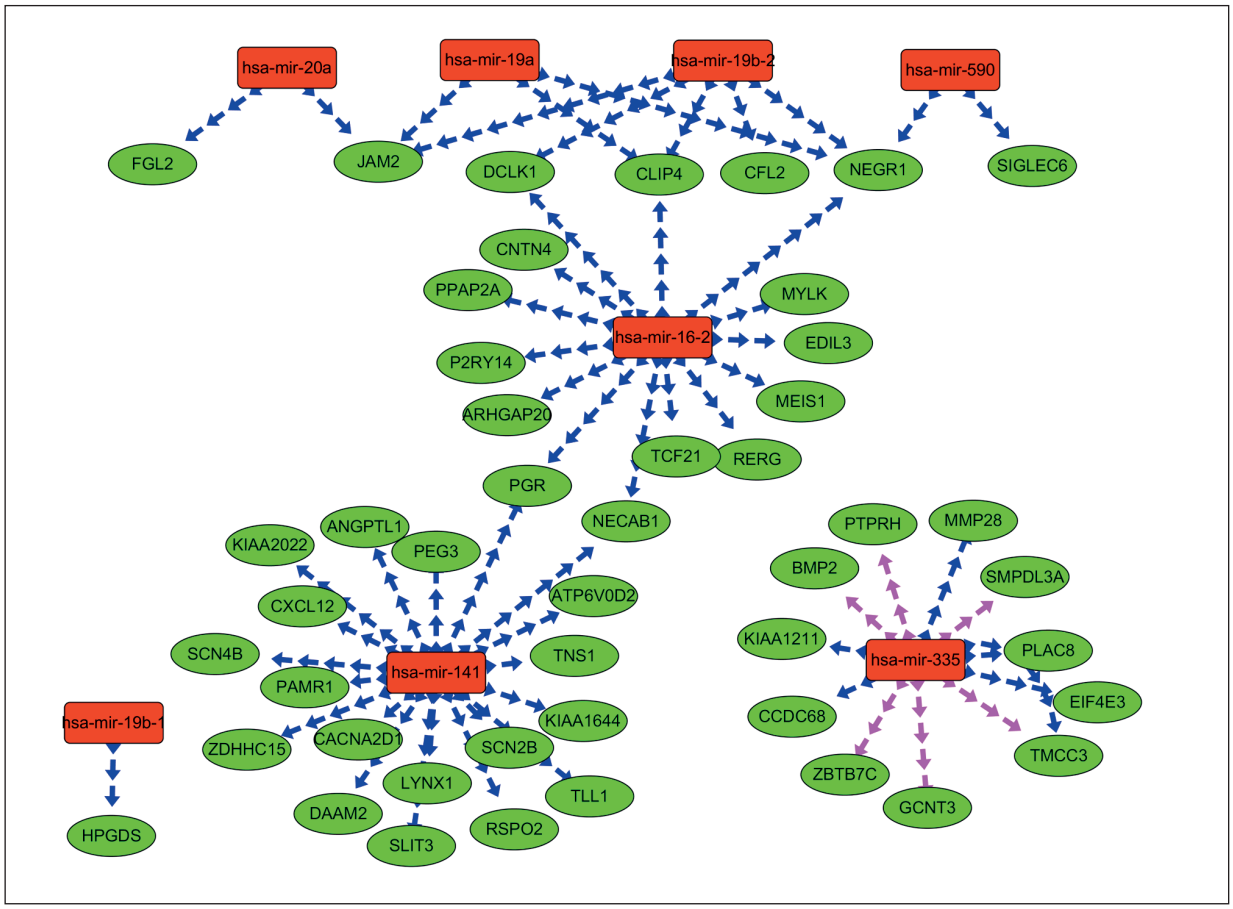

Figure 3. The miRNAs-targets interactions in COAD that were not only predicted by algorithms but also had the negative correlations. The green rectangle represented down-regulation of miRNAs, and the red ellipse represented the up-regulation of target genes. The arrow lines indicated miRNAs-targets pairs with negative correlations, in which blue color indicated the miRNAs-targets pairs were predicted by algorithms and pink color indicated the miRNAs-targets pairs were validated by experiment.

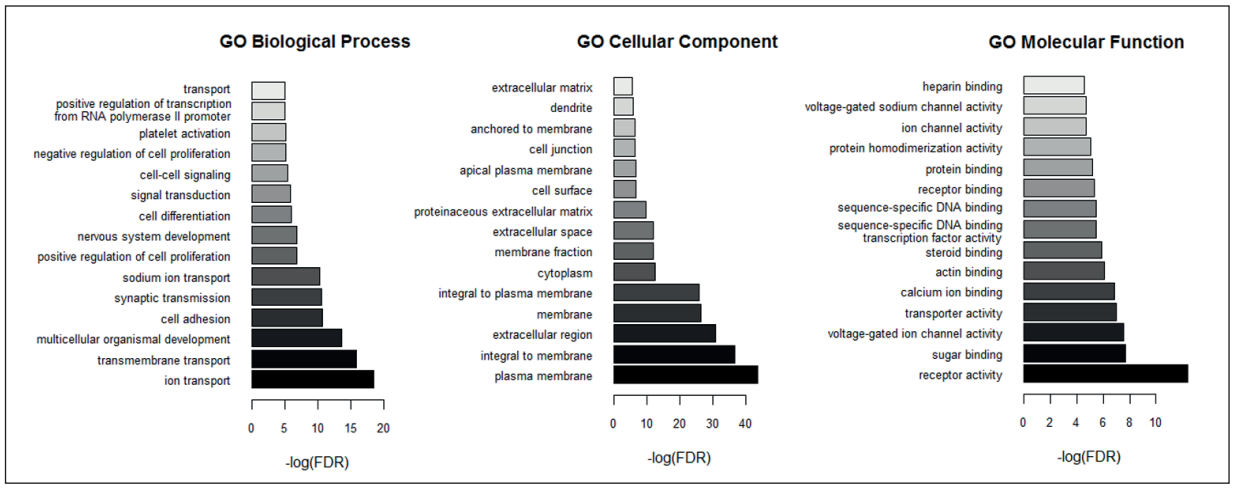

Figure 4. The significantly enriched GO terms of differentially miRNA targets in COAD compared to adjacent tissues.

Table II. Top 20 enriched KEGG pathways of differentially expressed target genes.

KEGG ID	Items	*p*-value	FDR
hsa04020	Calcium signaling pathway	4.77E-11	8.53E-09
hsa04724	Glutamatergic synapse	1.95E-09	1.17E-07
hsa04080	Neuroactive ligand-receptor interaction	1.47E-09	1.31E-07
hsa04060	Cytokine-cytokine receptor interaction	1.68E-08	7.51E-07
hsa04972	Pancreatic secretion	3.04E-08	1.09E-06
hsa04974	Protein digestion and absorption	1.27E-07	3.79E-06
hsa04270	Vascular smooth muscle contraction	1.22E-06	3.11E-05
hsa04970	Salivary secretion	2.49E-06	5.58E-05
hsa04960	Aldosterone-regulated sodium reabsorption	3.08E-06	6.13E-05
hsa04971	Gastric acid secretion	3.72E-06	6.66E-05
hsa04978	Mineral absorption	1.40E-05	0.000227
hsa00982	Drug metabolism - cytochrome P450	2.15E-05	0.000297
hsa03320	PPAR signaling pathway	2.15E-05	0.000297
hsa00512	Mucin type O-Glycan biosynthesis	4.06E-05	0.00052
hsa04062	Chemokine signaling pathway	6.58E-05	0.000785
hsa04640	Hematopoietic cell lineage	8.46E-05	0.000947
hsa04964	Proximal tubule bicarbonate reclamation	9.54E-05	0.001004
hsa00500	Starch and sucrose metabolism	0.000117	0.001159
hsa04976	Bile secretion	0.000159	0.001499
hsa04975	Fat digestion and absorption	0.000287	0.002572

Correlation of DNA Methylation and mRNA Expression

The DNA methylation was performed on the MethylJHU-USC Human Methylation450 platform, and a final cohort was consisted of 277 tumors and 19 adjacent normal tissues of the same patients. A total of 275 differentially methylated regions involved 235 genes were identified with FDR<0.05 in COAD.

The correlations between DNA methylation and mRNA expression were analyzed. We identified 11 genes with inverse correlations between mRNA expression and DNA methylation. All these 11 differentially mRNAs were down-regulated and hypermethylated in COAD compared with adjacent normal tissues including PDX1, SIM2, SLC4A11, MSX1, ACAN, TLX1, KRT7, TBX15, F7, IGF2, DLX5 (Table III).

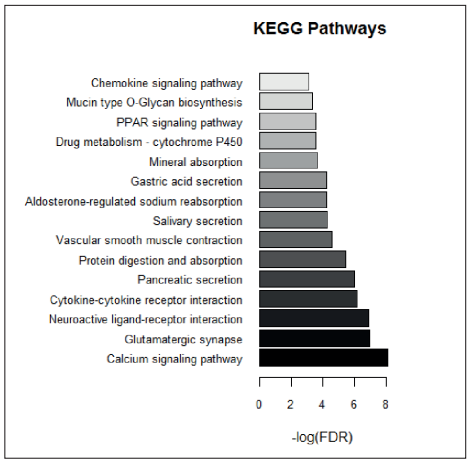

Figure 5. The significantly enriched KEGG pathways of differentially miRNA targets in COAD compared to adjacent tissues.

Confirmation of Differentially Expressed miRNAs and mRNAs

Three pairs of COAD and the adjacent tissues were used to verify the expression of integrated analysis. Two miRNAs (miR-141 and miR-335) and three targets (RSPO2, CXCL12 and BMP2) were selected. The results revealed that the expression of miR-141 and miR-335 showed significant up-regulation. Moreover, RSPO2 and CXCL12 were down-regulated, while BMP2 was up-regulated in COAD compared with the adjacent tissues (Figure 6).

According to GSE83924, miR141 (*p*-value=0.05755), miR19a (*p*-value=0.01209), miR20a (*p*-value=5.574E-07) was up-regulated in COAD. However, there was no significant difference in the expression of miR-19b-2 between tumors and adjacent normal tissues of COAD (Figure 7). The ROC curve of miR-20a between COAD patients and healthy controls was illustrated in Figure 8, suggesting its potential use in clinical diagnosis of COAD.

Table III. Differentially genes whose expression was inversely correlated with DNA methylation in COAD compared with adjacent tissues.

Gene_ID	Gene symbol	Log2 fold change	Position in cpg.island	FDR
6493	SIM2	-4.00027699	chr21:38079941-38081833	3.42834E-07
3651	PDX1	-4.792034073	chr13:28498226-28499046	7.90262E-06
1749	DLX5	-2.038160709	chr7:96650221-96651551	2.2815E-05
3481	IGF2	-2.834876417	chr11:2158951-2162484	8.80473E-05
4487	MSX1	-3.071848907	chr4:4861227-4862241	0.005017089
83959	SLC4A11	-4.481436767	chr20:3218578-3220930	0.005444961
176	ACAN	-2.854507436	chr15:89346043-89347203	0.009866599
3195	TLX1	-3.64693574	chr10:102893660-102895059	0.011652729
6913	TBX15	-2.895168069	chr1:119522385-119522638	0.018428163
3855	KRT7	-2.673534548	chr12:52626793-52627577	0.018495013
2155	F7	-2.83236385	chr13:113761566-113765534	0.026554773

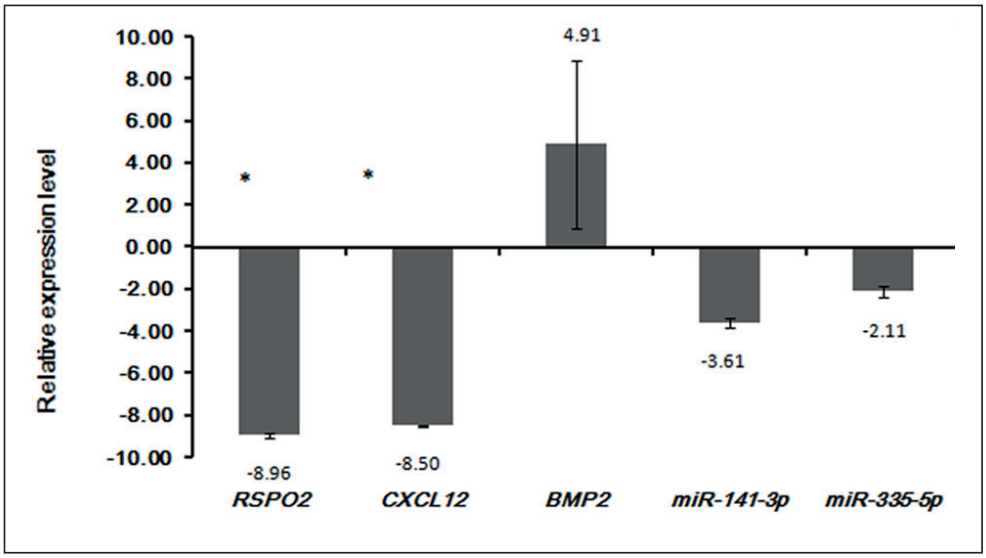

Figure 6. qRT-PCR results of RSPO2, CXCL12, BMP2, miR-141-3p and miR-335-5p in COAD. * $p<0.05$.

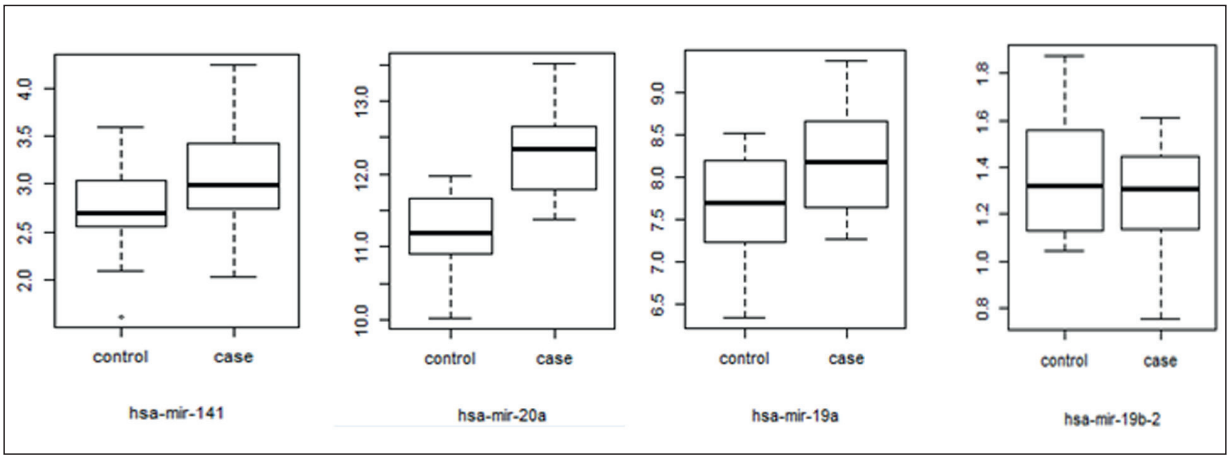

Figure 7. The expression of miR-141, miR-20a, miR-19a and miR-19b-2 in COAD based on GSE83924.

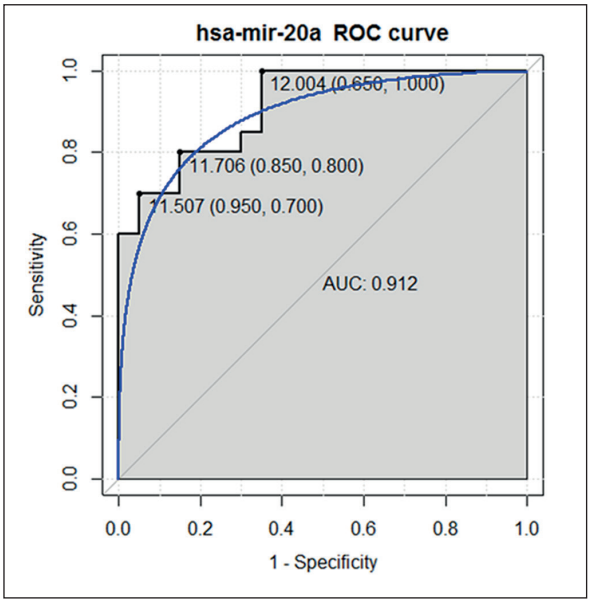

Figure 8. The ROC curve of miR-20a between COAD patients and healthy controls. The ROC curve is shown to evaluate diagnostic ability of miR-20a with the proportion of true positive (Sensitivity) and false positive (1-Specificity). With the optimal cutoff value, the proportion of Sensitivity is 85.0% and the proportion of 1-Specificity is 80.0%. The AUC (area under curve) value is 0.912 which indicates an accurately diagnostic ability.

Discussion

COAD remains the major type of colon cancer worldwide. Sensitive biomarkers can facilitate disease detection, staging and prediction of therapeutic outcome. To discover the non-invasive markers that can distinguish COAD from the surrounding normal tissues, we integrated the miRNA and mRNA expression data and methylation data from TCGA data portal to identify the entire transcription profile of COAD compared to the normal colon mucosa.

A previous study showed that plasma miR-141 was significantly associated with stage IV colon cancer and distant metastasis, and high levels of miR-141 in plasma were associated with poor prognosis[35]. Serum miR-141 was up-regulated in synchronous liver-metastatic colorectal cancer compared to consecutive localized colorectal cancer, and it was considered as a biomarker for early detection of liver metastasis in colorectal cancer[36,37]. A recent study also reported that miR-141 showed a trend of being higher in stage IV colorectal cancer (CRC) and can differentiate stage IV colorectal cancer from stage I-III patients[38]. The miRNA signature of stage III colorectal cancer showed that miR-141 was significantly over-expressed in tumor samples when compared with normal samples[39]. Induction of epithelial-mesenchymal transition and down-regulation of miR-141 was associated with oxaliplatin-resistant colorectal cancer cells[40]. RSPO2, one of the miRNA-141 targets, functioned as a tumour suppressor and was down-regulated in human CRCs due to promoter hypermethylation, which correlated with tumor differentiation, size and metastasis[41]. Comparing with normal tissues, we also found that miR-141 was up-regulated while RSPO2 was down-regulated in COAD. Moreover, miR-141 regulated most of the targets, indicating that miR-141 will be a promising biomarker for early detection of COAD.

It has been found that miR-19a can be induced by phosphatase of regenerating liver-3 and contribute to the increased proliferation and invasiveness of the colon cancer cells[42]. Huang et al[43] reported that high expression of miR-19a was associated with lymph node metastasis and played an important role in TNF-α-induced EMT in colorectal cancer cells and it may serve as a potential marker of lymph node metastasis. MiR-19a can be used to screen for patients with colorectal polyps, adenomas, or both[44]. Herein, miR-19a was up-regulated in COAD, showing the diagnostic value of serum expression of miR-19a in COAD patients.

MiR-19b has found to be associated with colon cancer by regulating the cell proliferation and cell cycle[45]. Moreover, up-regulation of miR-19b was detected in the transitional zone from normal to adenomatous tissue[46]. According to our miRNA and mRNA expression analysis, both miR-19b-1 and miR-19b-2 were up-regulated in tumor tissues of COAD which emphasize the importance of miR-19b in COAD. However, the levels of miR19b-2 made no difference between tumor and adjacent tissues of COAD based on GSE83924, which needs further research.

MiR-335 was reported up-regulated in colonic cancer tissues, compared with the para-cancerous control[47], and high miR-335 expression level was markedly associated with the tumor size and differentiation of CRC[48]. Increased expression of miR-335 was associated with clinical progression of CRC[49]. BMP2, one target gene for miR-335, was a secreted member of the BMP superfamily, which can inhibit normal and cancerous colonic epithelial cell growth and promotes apoptosis and differentiation[50,51]. Our findings showed that miR-335 was up-regulated in COAD, providing additional evidence that miR-335 could play an

important role in the carcinogenesis of COAD.

MiR-20a was reported to be an important miRNA that appeared to differentiate between carcinoma and normal mucosa for both colon and rectal cancer[52]. Increased expression of miR-20a in CRC tissues was associated with tumor invasion and lymph node metastasis by modulating matrix metalloproteinases[53]. MiR-20a was significantly higher in faecal samples from CRC patients, and miR-20a can be used as a potential non-invasive biomarker for CRC screening[54]. Moreover, miR-20a-5p was correlated with survival in CRC[55]. Our study suggested that the up-regulation of the miR-20a in tumors was significant, and the ROC curve of miR-20a revealed its application value in clinical diagnosis for COAD. The findings in our study highlighted the miR-20a as promising biomarkers for the screening and monitoring of COAD patients.

In our work, miR-590 was up-regulated in COAD according to the integrated analysis. Up-regulated miR-590 has been detected in a variety of cancers[56,57], suggesting that miR-590 may be a possible oncomiR. A recent paper[58] indicated that miR-590 can repress the angiogenesis and metastasis of colorectal cancer which may be a key regulator of the process of COAD.

MiR-16 has reported to be a tumor inhibitor in many types of cancers[59], miR-16 was down-regulated in 67% of colorectal cancer tissues and can repress the proliferation and induce apoptosis of tumors cells in colorectal cancer[59,60]. According to GSE83924, miR-16 was also down-regulated in tumor tissues of COAD which expressed in contrary pattern with that in our integrated analysis. The precise role of miR-16 needs further research.

Previous studies[61-65] have identified some key miRNAs associated with colorectal cancer through integrated analysis of paired miRNA-mRNA expression profiles based on GEO. In our study, miRNA-mRNA expression correlations analysis in COAD was performed based on TCGA and the results were different from that based on GEO, which provided new clues for the pathogenesis of COAD.

To understand the functional information of differentially expressed miRNAs, the miRNA targets were obtained by prediction or correlation analysis. Functional enrichment analysis of miRNA targets revealed that pancreatic secretion, salivary secretion, gastric acid secretion and bile secretion were significantly enriched, which were all associated with digestive tract cancer. A recent study[66] indicated that miRNAs will be potential cancer biomarker for digestive tract cancers. Our results proposed that the specific miRNA signature may be an emerging cancer biomarker for digestive tract cancers.

To identify sensitive and specific DNA methylation markers of COAD, we analyzed the DNA methylation of COAD, and further investigated the influence of methylation on gene expression. Compared to adjacent tissues, a total of 11 down-regulated and high-melthylated genes (PDX1, SIM2, SLC4A11, MSX1, ACAN, TLX1, KRT7, TBX15, F7, IGF2, and DLX5) in COAD were identified. The expression of these 11 genes may be inhibited by high-methylation.

Among those genes, MSX1 is a member of the muscle segment homeobox gene family. Deficiency of MSX1 was common in a variety of cancers and over-expressed MXS1 can repress the proliferation of tumor cells[67-69]. In addition, MXS1 has found to be methylated in breast cancer, colon cancer, lung cancer and prostate cancer frequently which suggested that methylation of MSX1 may be a tumor-acquired methylation. Hence, hypermethylation of MXS1 was speculated to play a role in COAD by down-regulation of MXS1.

KRT7 (cytokeratin 7) has been served as an effective immunohistochemical diagnostic tool to study cancer-initiating mechanisms in multiple organs[70]. The KRT7-/KRT20+ expression pattern has found to be highly traits of colorectal cancer[71-73]. Previous studies[74,75] have reported that negative KRT7 is typically detected in adenocarcinoma of the colorectal region. We speculated that hypermeltylation of KRT7 may involve in the pathogenesis of COAD.

Conclusions

Our findings identified eight differentially miRNAs (hsa-mir-141, hsa-mir-16-2, hsa-mir-19a, hsa-mir-19b-1, hsa-mir-19b-2, hsa-mir-20a, hsa-mir-335 and hsa-mir-590) in patients with COAD compared to normal colon tissues which may be potential diagnostic biomarkers. Together, miRNA seems to be an excellent tool for effectively monitoring and targeting COAD. Two differentially methylated and expressed genes (MSX1 and KRT7) in COAD were identified which may provide new clues for the strategies of diagnose and treatment.

Conflict of Interest

The Authors declare that they have no conflict of interests.

SUPPLEMENTARY FIGURES

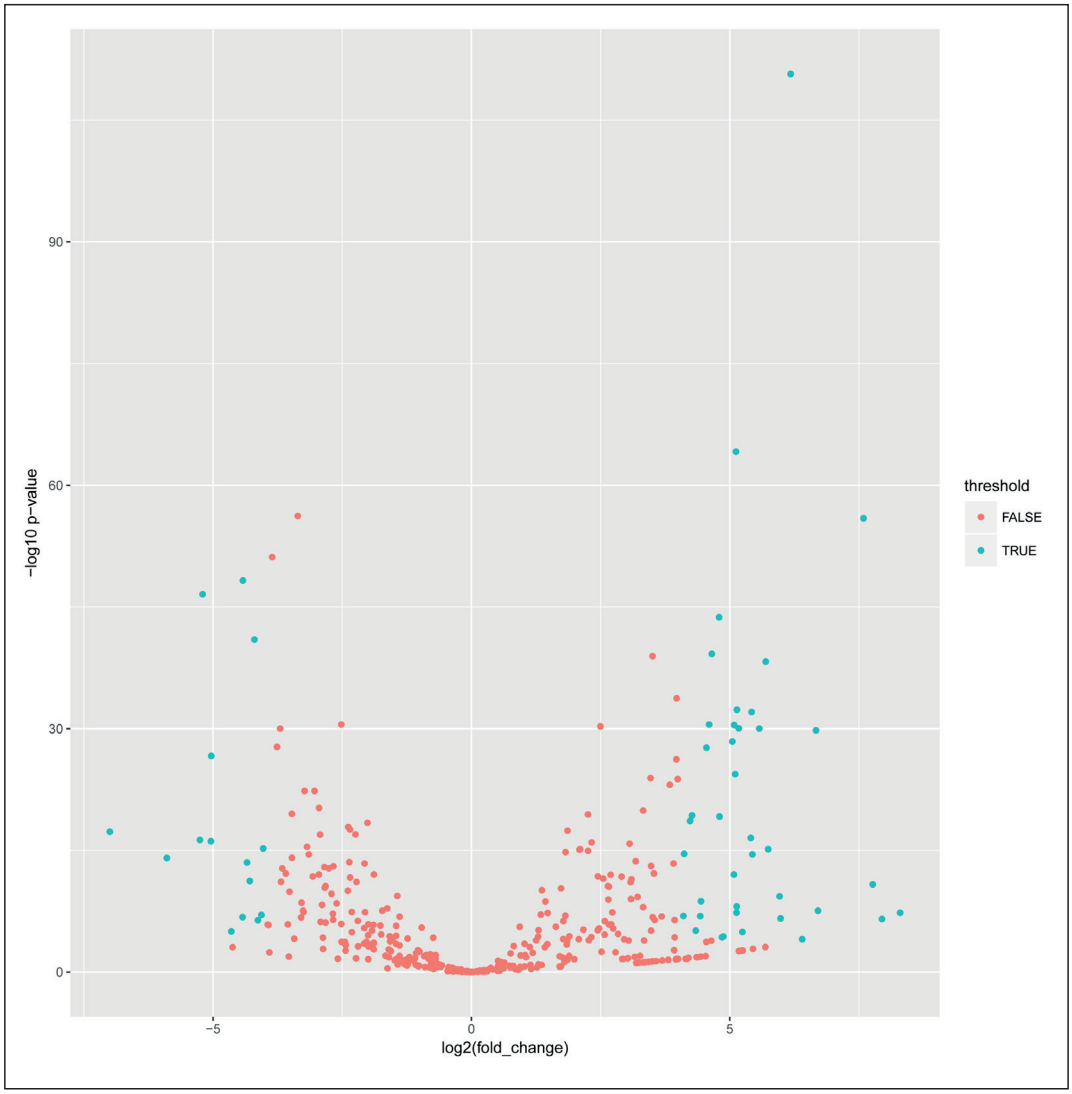

Figure S1. The volcano plots of differentially expressed miRNAs.

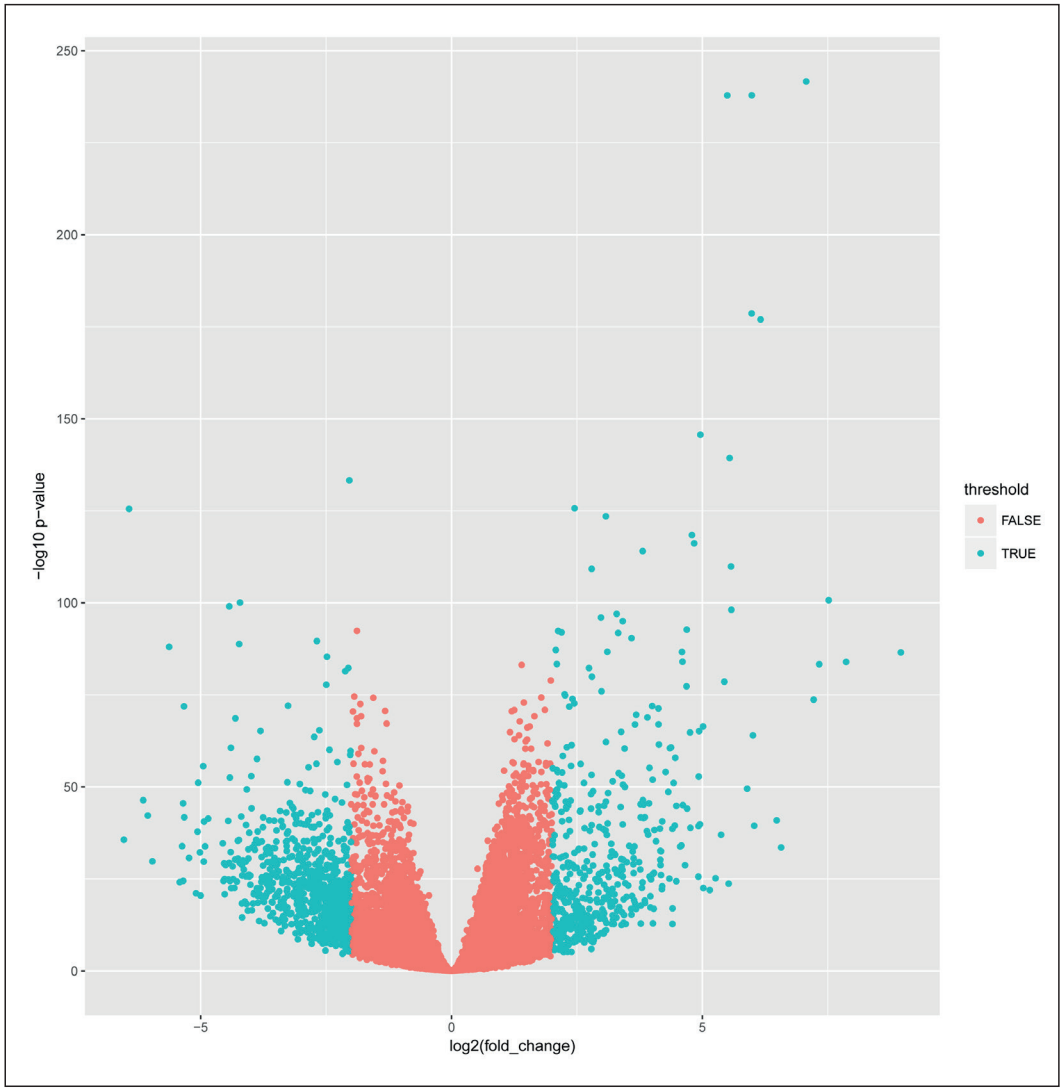

Figure S2. The volcano plots of differentially expressed mRNAs.

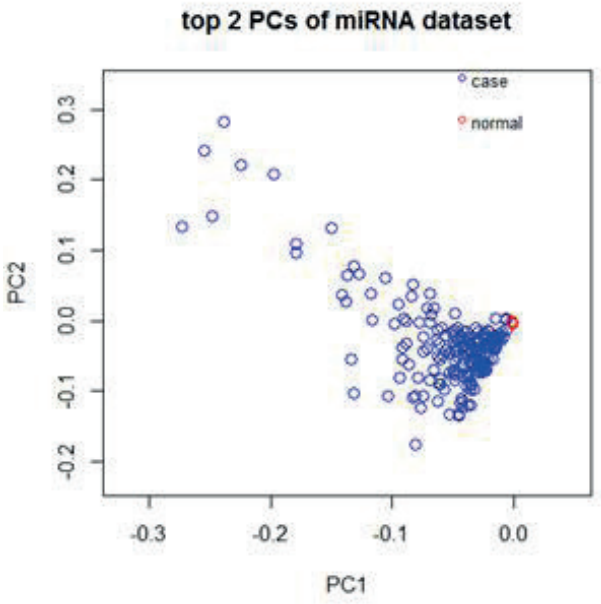

Figure S3. The principal component analysis of miRNA expression data (case=202 and normal control=5).

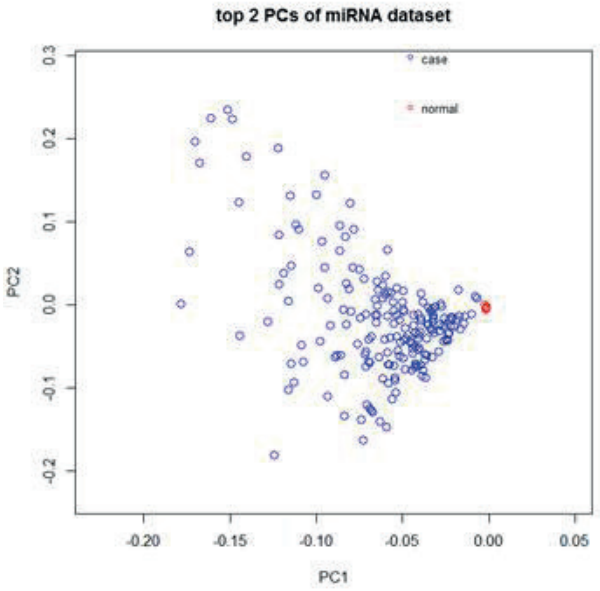

Figure S4. The principal component analysis of miRNA expression data after 9 discrete cases were eliminated (case=193 and normal control=5).

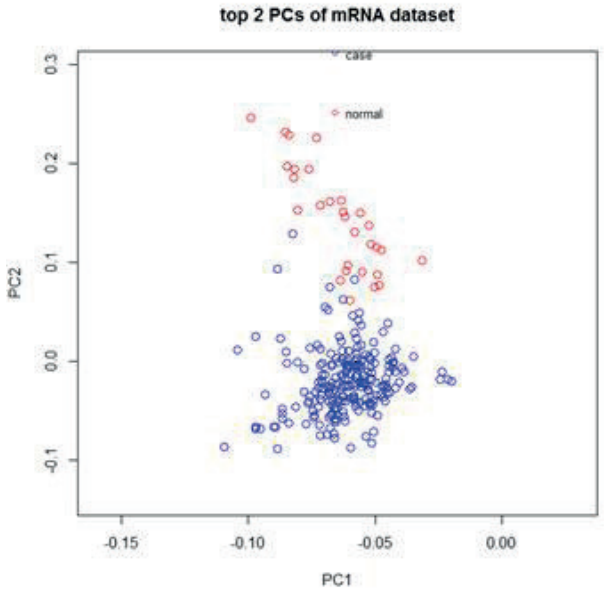

Figure S5. The principal component analysis of mRNA expression data (case=219 and normal control=29).

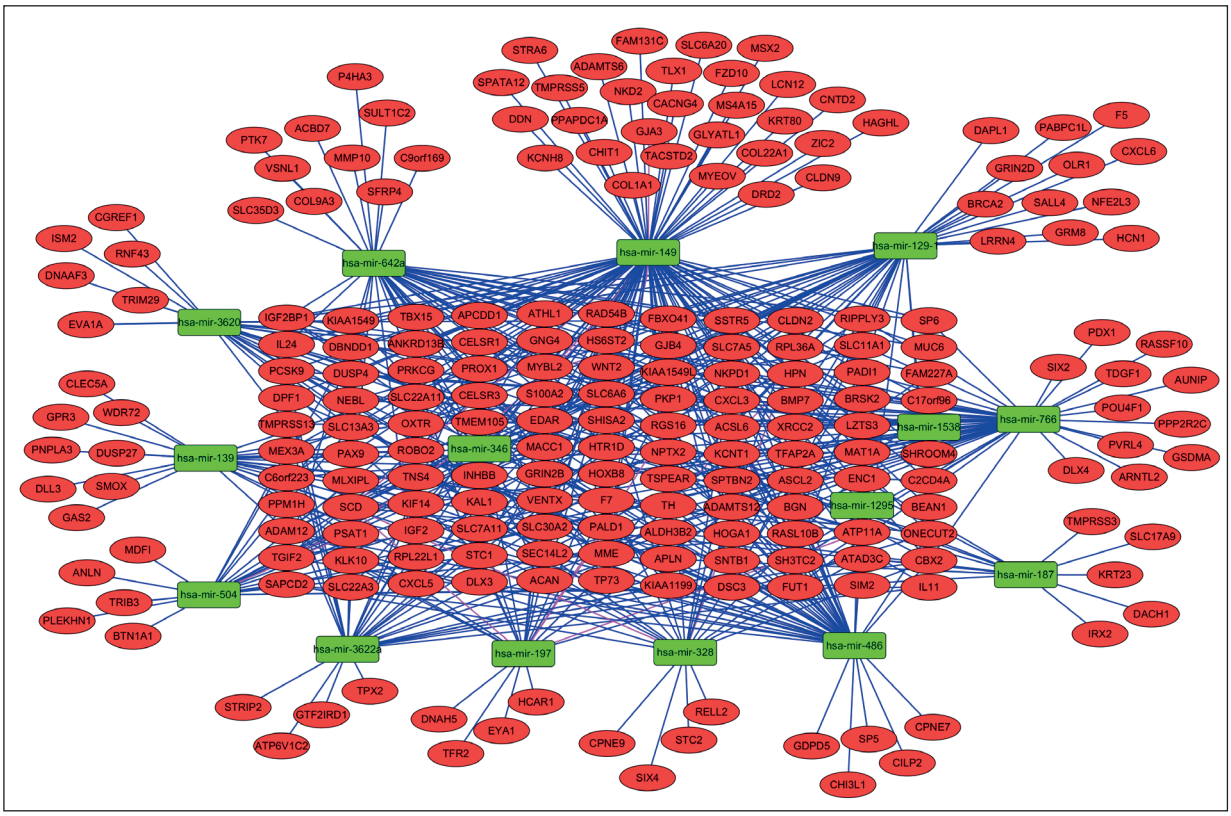

Figure S6. The down-regulated miRNA-mRNA pairs in COAD compare to adjacent tissues.

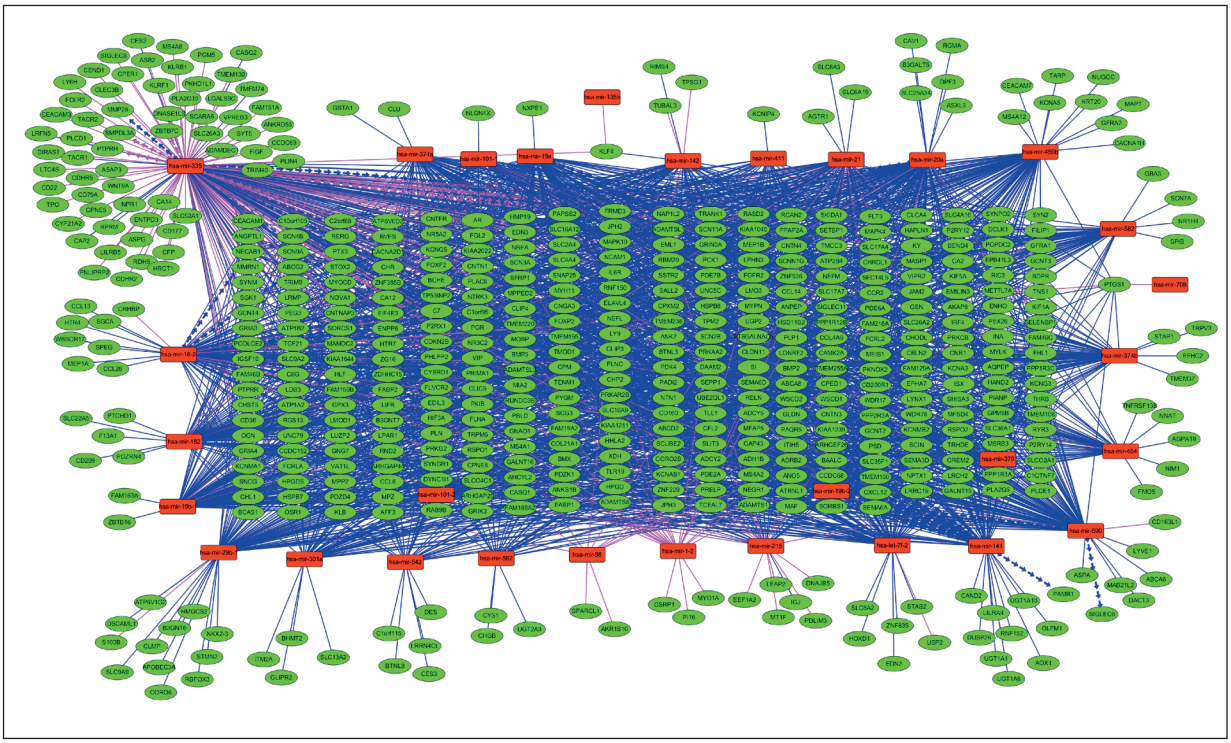

Figure S7. The up-regulated miRNA-mRNA pairs in COAD compare to adjacent tissues.

References

1) Willaert W, Ceelen W. Extent of surgery in cancer of the colon: is more better? World J Gastroenterol 2015; 21: 132-138.
2) Bae JA, Kho DH, Sun EG, Ko YS, Yoon S, Lee KH, Ahn KY, Lee JH, Joo YE, Chung IJ, Lee SH, Kim H, Kim KK. Elevated coexpression of KITENIN and the ErbB4 CYT-2 isoform promotes the transition from colon adenoma to carcinoma following APC loss. Clin Cancer Res 2016; 22: 1284-1294.
3) Sewda K, Coppola D, Enkemann S, Yue B, Kim J, Lopez AS, Wojtkowiak JW, Stark VE, Morse B, Shibata D, Vignesh S, Morse DL. Cell-surface markers for colon adenoma and adenocarcinoma. Oncotarget 2016; 7: 17773-17789.
4) Missiaglia E, Jacobs B, D'Ario G, Di Narzo AF, Soneson C, Budinska E, Popovici V, Vecchione L, Gerster S, Yan P, Roth AD, Klingbiel D, Bosman FT, Delorenzi M, Tejpar S. Distal and proximal colon cancers differ in terms of molecular, pathological, and clinical features. Ann Oncol 2014; 25: 1995-2001.
5) Mohney BG, Robertson DM, Schomberg PJ, Hodge DO. Second nonocular tumors in survivors of heritable retinoblastoma and prior radiation therapy. Am J Ophthalmol 1998; 126: 269-277.
6) Grone J, Lenze D, Jurinovic V, Hummel M, Seidel H, Leder G, Beckmann G, Sommer A, Grutzmann R, Pilarsky C, Mansmann U, Buhr HJ, Stein H, Hummel M. Molecular profiles and clinical outcome of stage UICC II colon cancer patients. Int J Colorectal Dis 2011; 26: 847-858.
7) Yusuf F, Ilyas S, Damanik HA, Fatchiyah F. Microbiota composition, HSP70 and Caspase-3 expression as marker for colorectal cancer patients in Aceh, Indonesia. Acta Med Indones 2016; 48: 289-299.
8) Simsek O, Kocael A, Kemik A, Vatankulu B, Kocael P, Halac M, Ulualp K, Sonmezoglu K. Decreased preoperative serum fibulin-3 levels in colon cancer patients. Eur Rev Med Pharmacol Sci 2015; 19: 4076-4080.
9) Uddin S, Ahmed M, Hussain A, Abubaker J, Al-Sanea N, AbdulJabbar A, Ashari LH, Alhomoud S, Al-Dayel F, Jehan Z, Bavi P, Siraj AK, Al-Kuraya KS. Genome-wide expression analysis of Middle Eastern colorectal cancer reveals FOXM1 as a novel target for cancer therapy. Am J Pathol 2011; 178: 537-547.
10) Hong Y, Ho KS, Eu KW, Cheah PY. A susceptibility gene set for early onset colorectal cancer that integrates diverse signaling pathways: implication for tumorigenesis. Clin Cancer Res 2007; 13: 1107-1114.
11) Zhang HQ, Guo X, Guo SQ, Wang Q, Chen XQ, Li XN, Guo LS. STMN1 in colon cancer: expression and prognosis in Chinese patients. Eur Rev Med Pharmacol Sci 2016; 20: 2038-2044.
12) Cai X, Qi WX, Wang L, Zhang Z. Correlation of multiple proteins with clinic-pathological features and its prognostic significance in colorectal cancer with signet-ring cell component. Eur Rev Med Pharmacol Sci 2016; 20: 3358-3367.
13) Lin G, He X, Ji H, Shi L, Davis RW, Zhong S. Reproducibility Probability Score--incorporating measurement variability across laboratories for gene selection. Nat Biotechnol 2006; 24: 1476-1477.
14) Hu Q, Wang YB, Zeng P, Yan GQ, Xin L, Hu XY. Expression of long non-coding RNA (lncRNA) H19 in immunodeficient mice induced with human colon cancer cells. Eur Rev Med Pharmacol Sci 2016; 20: 4880-4884.
15) Schetter AJ, Okayama H, Harris CC. The role of microRNAs in colorectal cancer. Cancer J 2012; 18: 244-252.
16) Faber C, Kirchner T, Hlubek F. The impact of microRNAs on colorectal cancer. Virchows Arch 2009; 454: 359-367.
17) Slaby O, Svoboda M, Michalek J, Vyzula R. MicroRNAs in colorectal cancer: translation of molecular biology into clinical application. Mol Cancer 2009; 8: 102.
18) Heyn H, Esteller M. DNA methylation profiling in the clinic: applications and challenges. Nat Rev Genet 2012; 13: 679-692.
19) Esteller M. Epigenetics in cancer. N Engl J Med 2008; 358: 1148-1159.
20) Esteller M. Epigenetic gene silencing in cancer: the DNA hypermethylome. Human Mol Genet 2007; 16: R50-R59.
21) Robertson KD. DNA methylation and human disease. Nat Rev Genet 2005; 6: 597-610.
22) Wang Y, Zhang J, Li L, Xu X, Zhang Y, Teng Z, Wu F. Identification of molecular targets for predicting colon adenocarcinoma. Med Sci Monit 2016; 22: 460-468.
23) Esteller M. Epigenetic gene silencing in cancer: the DNA hypermethylome. Hum Mol Genet 2007; 16 Spec No 1: R50-59.
24) Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biol 2010; 11: 1-12.
25) Benjamini Y, Hochberg Y. Controlling the false discovery rate--a practical and powerful approach to multiple testing. J R Stat Soc Ser A 1995; 57: 289-300.
26) Singh M, Simpson SM, Mullins HR, Dickson JS. Thermal tolerance of acid-adapted and non-adapted Escherichia coli O157:H7 and Salmonella in ground beef during storage. Foodborne Pathog Dis 2006; 3: 48-61.
27) Tabas-Madrid D, Nogales-Cadenas R, Pascual-Montano A. GeneCodis3: a non-redundant and modular enrichment analysis tool for functional genomics. Nucleic Acids Res 2012; 40: W478-483.
28) Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000; 25: 25-29.
29) Kanehisa M, Goto S. Kegg. Kyoto Encyclopedia of Genes and Genomes. Nucleic Acids Res 2000; 28: 27-30.
30) Warden CD, Lee H, Tompkins JD, Li X, Wang C, Riggs AD, Yu H, Jove R, Yuan YC. COHCAP: an integrative genomic pipeline for single-nucleotide resolution DNA methylation analysis. Nucleic Acids Res 2013; 41: e117-e117.
31) Wang Y, Zhang J, Li L, Xu X, Zhang Y, Teng Z, Wu F. Identification of molecular targets for predicting colon adenocarcinoma. Med Sci Monit 2016; 22: 460-468.

32) Cheng H, Zhang L, Cogdell DE, Zheng H, Schetter AJ, Nykter M, Harris CC, Chen K, Hamilton SR, Zhang W. Circulating plasma MiR-141 is a novel biomarker for metastatic colon cancer and predicts poor prognosis. PLoS One 2011; 6: e17745.
33) Yin J, Bai Z, Song J, Yang Y, Wang J, Han W, Zhang J, Meng H, Ma X, Yang Y, Wang T, Li W, Zhang Z. Differential expression of serum miR-126, miR-141 and miR-21 as novel biomarkers for early detection of liver metastasis in colorectal cancer. Chin J Cancer Res 2014; 26: 95-103.
34) Yang Y, Gu X, Zhou M, Xiang J, Chen Z. Serum microRNAs: a new diagnostic method for colorectal cancer. Biomed Rep 2013; 1: 495-498.
35) Sun Y, Liu Y, Cogdell D, Calin GA, Sun B, Kopetz S, Hamilton SR, Zhang W. Examining plasma microRNA markers for colorectal cancer at different stages. Oncotarget 2016 7: 11434-11449.
36) Brunet Vega A, Pericay C, Moya I, Ferrer A, Dotor E, Pisa A, Casalots A, Serra-Aracil X, Oliva JC, Ruiz A, Saigi E. microRNA expression profile in stage III colorectal cancer: circulating miR-18a and miR-29a as promising biomarkers. Oncol Rep 2013; 30: 320-326.
37) Tanaka S, Hosokawa M, Yonezawa T, Hayashi W, Ueda K, Iwakawa S. Induction of epithelial-mesenchymal transition and down-regulation of miR-200c and miR-141 in oxaliplatin-resistant colorectal cancer cells. Biol Pharm Bull 2015; 38: 435-440.
38) Wu C, Qiu S, Lu L, Zou J, Li WF, Wang O, Zhao H, Wang H, Tang J, Chen L, Xu T, Sun Z, Liao W, Luo G, Lu X. RSPO2-LGR5 signaling has tumour-suppressive activity in colorectal cancer. Nat Commun 2014; 5: 3149.
39) Zhang J, Xiao Z, Lai D, Sun J, He C, Chu Z, Ye H, Chen S, Wang J. miR-21, miR-17 and miR-19a induced by phosphatase of regenerating liver-3 promote the proliferation and metastasis of colon cancer. Br J Cancer 2012; 107: 352-359.
40) Huang L, Wang X, Wen C, Yang X, Song M, Chen J, Wang C, Zhang B, Wang L, Iwamoto A, Wang J, Liu H. Hsa-miR-19a is associated with lymph metastasis and mediates the TNF-alpha induced epithelial-to-mesenchymal transition in colorectal cancer. Sci Rep 2015; 5: 13350.
41) Verma AM, Patel M, Aslam MI, Jameson J, Pringle JH, Wurm P, Singh B. Circulating plasma microRNAs as a screening method for detection of colorectal adenomas. Lancet 2015; 385 Suppl 1: S100.
42) Kurokawa K, Tanahashi T, Iima T, Yamamoto Y, Akaike Y, Nishida K, Masuda K, Kuwano Y, Murakami Y, Fukushima M, Rokutan K. Role of miR-19b and its target mRNAs in 5-fluorouracil resistance in colon cancer cells. J Gastroenterol 2012; 47: 883-895.
43) Knudsen KN, Nielsen BS, Lindebjerg J, Hansen TF, Holst R, Sørensen FB. microRNA-17 is the most up-regulated member of the miR-17-92 cluster during early colon cancer evolution. PLoS One 2015; 10.
44) Wang YX, Zhang XY, Zhang BF, Yang CQ, Chen XM, Gao HJ. Initial study of microRNA expression profiles of colonic cancer without lymph node metastasis. J Dig Dis 2010; 11: 50-54.
45) Lu Y, Yang H, Yuan L, Liu G, Zhang C, Hong M, Liu Y, Zhou M, Chen F, Li X. Overexpression of miR-335 confers cell proliferation and tumour growth to colorectal carcinoma cells. Mol Cell Biochem 2016; 412: 235-245.
46) Vickers MM, Bar J, Gorn-Hondermann I, Yarom N, Daneshmand M, Hanson JE, Addison CL, Asmis TR, Jonker DJ, Maroun J, Lorimer IA, Goss GD, Dimitroulakos J. Stage-dependent differential expression of microRNAs in colorectal cancer: potential role as markers of metastatic disease. Clin Exp Metastasis 2012; 29: 123-132.
47) Hardwick JC, Van Den Brink GR, Bleuming SA, Ballester I, Van Den Brande JM, Keller JJ, Offerhaus GJ, Van Deventer SJ, Peppelenbosch MP. Bone morphogenetic protein 2 is expressed by, and acts upon, mature epithelial cells in the colon. Gastroenterology 2004; 126: 111-121.
48) Beck SE, Jung BH, Fiorino A, Gomez J, Rosario ED, Cabrera BL, Huang SC, Chow JY, Carethers JM. Bone morphogenetic protein signaling and growth suppression in colon cancer. Am J Physiol Gastrointest Liver Physiol 2006; 291: G135-145.
49) Pellatt DF, Stevens JR, Wolff RK, Mullany LE, Herrick JS, Samowitz W, Slattery ML. Expression profiles of miRNA subsets distinguish human colorectal carcinoma and normal colonic mucosa. Clin Transl Gastroenterol 2016; 7: e152.
50) Xu T, Jing C, Shi Y, Miao R, Peng L, Kong S, Ma Y, Li L. microRNA-20a enhances the epithelial-to-mesenchymal transition of colorectal cancer cells by modulating matrix metalloproteinases. Exp Ther Med 2015; 10: 683-688.
51) Yau TO, Wu CW, Tang CM, Chen Y, Fang J, Dong Y, Liang Q, Ng SS, Chan FK, Sung JJ, Yu J. microRNA-20a in human faeces as a non-invasive biomarker for colorectal cancer. Oncotarget 2016; 7: 1559-1568.
52) Chen X, Shi K, Wang Y, Song M, Zhou W, Tu H, Lin Z. Clinical value of integrated-signature miRNAs in colorectal cancer: miRNA expression profiling analysis and experimental validation. Oncotarget 2015; 6: 37544-37556.
53) Shen B, Yu S, Zhang Y, Yuan Y, Li X, Zhong J, Feng J. miR-590-5p regulates gastric cancer cell growth and chemosensitivity through RECK and the AKT/ERK pathway. Onco Targets Ther 2016; 9: 6009-6019.
54) Jiang X, Xiang G, Wang Y, Zhang L, Yang X, Cao L, Peng H, Xue P, Chen D. MicroRNA-590-5p regulates proliferation and invasion in human hepatocellular carcinoma cells by targeting TGF-beta RII. Mol Cells 2012; 33: 545-551.
55) Zhou Q, Zhu Y, Wei X, Zhou J, Liang C, Hong S, Yu H, Piao D, Sha R, Bai Y. MiR-590-5p inhibits colorectal cancer angiogenesis and metastasis by regulating nuclear factor 90l[sol]lvascular endothelial growth factor A axis. Cell Death Dis 2016; 7.
56) Ma Q, Wang X, Li Z, Li B, Ma F, Peng L, Zhang Y, Xu A, Jiang B. microRNA-16 represses colorectal cancer cell growth in vitro by regulating the p53/survivin signaling pathway. Oncol Rep 2013; 29: 1652-1658.
57) Kou Y, Qiao L, Wang Q. Identification of core miRNA based on small RNA-seq and RNA-seq for colorectal cancer by bioinformatics. Tumor Biol 2015; 36: 2249-2255.
58) Zhao X, Song H, Zuo Z, Zhu Y, Dong X, Lu X. Identification of miRNA-miRNA synergistic relationships in colorectal cancer. Int J Biol Macromol 2012; 55: 98-103.

59) Abba M, Benner A, Patil N, Heil O, Allgayer H. Differentially expressed microRNAs in colorectal cancer metastasis. Genom Data 2015; 6: 33-35.
60) Wang J, Yu H, Ye L, Jin L, Yu M, Lv Y. Integrated regulatory mechanisms of miRNAs and targeted genes involved in colorectal cancer. Int J Clin Exp Pathol 2015; 8: 517-529.
61) Pradhan M, Nagulapalli K, Ledford L, Pandit Y, Palakal M. A system biology approach for understanding the miRNA regulatory network in colon rectal cancer. Int J Data Min Bioinform 2015; 11.
62) Lu YC, Chang JT, Chan EC, Chao YK, Yeh TS, Chen JS, Cheng AJ. miR-196, an emerging cancer biomarker for digestive tract cancers. J Cancer 2016; 7: 650-655.
63) Park K, Kim K, Rho SB, Choi K, Kim D, Oh SH, Park J, Lee SH, Lee JH. Homeobox Msx1 interacts with p53 tumor suppressor and inhibits tumor growth by inducing apoptosis. Cancer Res 2005; 65: 749-757.
64) Park J, Park K, Kim S, Lee JH. Msx1 gene overexpression induces G1 phase cell arrest in human ovarian cancer cell line OVCAR3. Biochem Biophys Res Commun 2001; 281: 1234-1240.
65) Sliwinski T, Synowiec E, Czarny P, Gomulak P, Forma E, Morawiec Z, Morawiec J, Dziki L, Wasylecka M, Blasiak J. The c.469+46_56del mutation in the homeobox MSX1 gene--a novel risk factor in breast cancer? Cancer Epidemiol 2010; 34: 652-655.
66) Gao HW, Yu CP, Jin JS, Lee HS. Expression of keratin 7 in the boundary epithelial cells of colorectal neoplasm. J Med Sci 2008; 28.
67) Wang N. Coordinate expression of cytokeratin 7 and 20 defines unique subsets of carcinomas. Appl Immunohistochem 1995; 3: 99-107.
68) Chu P, Wu E, Weiss LM. Cytokeratin 7 and cytokeratin 20 expression in epithelial neoplasms: a survey of 435 cases. Mod Pathol 2000; 13: 962-972.
69) Loy TS, Calaluce RD, Keeney GL. Cytokeratin immunostaining in differentiating primary ovarian carcinoma from metastatic colonic adenocarcinoma. Mod Pathol 1996; 9: 1040-1044.
70) Chu PG, Weiss LM. Keratin expression in human tissues and neoplasms. Histopathology 2002; 41: 365-366.
71) Moll R. Cytokeratins as markers of differentiation in the diagnosis of epithelial tumors. Subcell Biochem 1998; 31: 205-262.