Identification of Key Genes in Colorectal Cancer Regulated by miR-34a

Background: The aim of this study was to screen the molecular targets of miR-34a in colorectal cancer (CRC) and construct the regulatory network, to gain more insights to the pathogenesis of CRC.

Material/Methods: The microarray data of CRC samples and normal samples (GSE4988), as well as CRC samples transformed with miR-34a and non-transfected CRC samples (GSE7754), were downloaded from the Gene Expression Omnibus (GEO) database. The differently expressed genes (DEGs) were identified via the LIMMA package in R language. The Database for Annotation, Visualization and Integrated Discovery (DAVID) was used to identify significant Gene Ontology (GO) terms and pathways in DEGs. The targets of miR-34a were obtained via the miRWalk database, and then the overlaps between them were selected out to construct the regulatory network of miR-34a in CRC using the Cytoscape software.

Results: A total of 392 DEGs were identified in CRC samples compared with normal samples, including 239 upregulated genes and 153 downregulated ones. These DEGs were enriched in 75 GO terms and one Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway. At the same time, 332 DEGs (188 upregulated and 144 downregulated) were screened out between miR-34a transformed CRC and miR-34a non-transfected CRC samples and they were enriched in 20 GO terms and eight KEGG pathways. Six overlapped genes were identified in two DEGs groups. There were 1,668 targets of miR-34a obtained via the miRWalk database, among which 21 were identified differently expressed in miR-34a transformed CRC samples compared with miR-34a non-transfected CRC samples. Two regulatory networks of miR-34a in CRC within these two groups of overlapped genes were constructed respectively.

Conclusions: Pathways related to cell cycle, DNA replication, oocyte meiosis, and pyrimidine metabolism might play critical roles in the progression of CRC. Several genes such as SERPINE1, KLF4, SEMA4B, PPARG, CDC45, and KIAA0101 might be the targets of miR-34a and the potential therapeutic targets of CRC.

MeSH Keywords: Colorectal Neoplasms • Microarray Analysis • MicroRNAs

Full-text PDF: https://www.medscimonit.com/abstract/index/idArt/904937
Background

Colorectal cancer (CRC) is the third most common cancer and the fourth-leading cause of cancer related death worldwide with an incidence of more than 1,000,000 per year [1,2]. According to statistics, the probability of developing CRC increases from 0.07% in the first four decades of life to 4.5–5% in the seventh decade of life [3,4]. Furthermore, studies suggest that 40–50% of patients who underwent potentially curative surgery ultimately relapsed and died of metastatic disease [5,6]. The most important prognostic indicator of survival in early CRC is the stage of the tumor determined by the depth of penetration through the bowel wall, and the number of involved lymph nodes [6]. However, the tumors are often diagnosed at an intermediate or late stage with poor prognosis, because of the complexity in medical diagnosis. In addition, pathological staging fails to predict recurrence accurately in many patients undergoing surgery for CRC. In fact, 10–20% patients with stage II CRC, and 30–40% of those with stage III CRC, develop recurrence [7]. It is crucial to explore the molecular mechanism and identify reliable biomarkers that can guide the diagnosis and therapy of CRC.

MicroRNAs (miRNAs) are a class of endogenous small non-coding RNAs to regulate gene expression [8], and they were reported to be closely associated with cell growth, proliferation, differentiation, and death [9,10]. There has been increasing evidence indicating that microRNAs play important roles in the development of cancers [11], such as the proliferation, apoptosis, invasion, and metastasis of tumors [12,13]. MicroRNA-34a (miR-34a) is a pivotal member of the p53 network, and it was found to be downregulated in multiple types of tumors, including CRC [14], and often act as a tumor suppressor [15–19]. Studies have shown that miR-34a regulated multiple developmental cell-fate mechanisms, including the differentiation of human embryonic stem cells and somatic cell reprogramming [16,20,21]. Furthermore, it has also been suggested that miR-34a plays critical roles in inhibiting tumor recurrence [22]. However, the specific regulatory mechanism of miR-34a in CRC is still unclear.

In this study, we aimed to explore some key targets associated with the development of CRC. And our study might contribute to promoting available biomarkers for the early diagnosis, therapy, and prognosis of CRC.

Material and Methods

Data preprocessing and identification of DEGs

The raw data was background corrected, log2 transformed, and quantile normalized using Robust Multi-array Average (RMA). If multiple probes corresponded to one gene, the mean expression value was defined as expression value. LIMMA package in R language was used to analyze the DEGs in CRC samples compared to normal samples (regarded as DEGs-1), as well as CRC samples transformed with miR-34a compared to non-transfected CRC samples (regarded as DEGs-2). DEGs were obtained according to the criteria: adjusted p<0.05 and |log(fold change)| >1.

Functional enrichment analysis of DEGs

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis were commonly used approaches for functional and pathway studies of large-scale genomic or transcription data, respectively [23]. GO terms included biological processes (BP), molecular function (MF), and cellular component (CC). The Database for Annotation, Visualization and Integrated Discovery (DAVID) [24] (https://david.ncifcrf.gov/) was a widely used web-based tool for the functional annotation of DEGs. To investigate the bio-functions of DEGs, GO, and KEGG pathway analyses were conducted based on the online software DAVID with p<0.05.

Construction of the transcriptional regulatory network of miR-34a

MiRWalk (http://mirwalk.uni-hd.de/) is a publicly available comprehensive resource, hosting the predicted as well as the experimentally validated microRNA (miRNA)-target interaction pairs. The targets of miR-34a were identified based on the miRWalk database, and the overlapped genes between these targets and DEGs-2 were selected. Then the transcriptional regulatory network with miR-34a was constructed. At the same time, the overlapped genes between DEGs-1 and DEGs-2 were also selected to build the regulatory network with miR-34a.

Verification of related upregulated and downregulated genes

The expression levels of some DEGs were detected in colon cancer cell line HCT116 (observed by our laboratory) by RT-PCR.
The culture of HCT116 and the retroviral expression of miR-34a referred to the methods of Chang et al. [25]. HCT116 cells transformed with miR-34a were named as miR-34a group, and non-transformed cells as the control group. Total RNA was isolated using TRIzol (Invitrogen, Carlsbad, CA, USA) according to the manufacturer’s instructions. Approximately 1 µg of total RNA from each sample was subjected to reverse transcription by SuperScript II reverse transcriptase (Invitrogen, America). Reverse transcription was performed at 42°C for one hour, followed by 95°C for five minutes. Quantitative real-time PCR (Q-PCR) was performed to determine the mRNA levels of SERPINE1, KLF4, SEMA4B, PPARG, CDC45, and KIAA0101 using a SYBR® Premix Ex Taq™ Kit (TaKaRa, Shiga, Japan) according to the manufacturer’s instructions and an ABI Prism 7000 Sequence Detection System (Applied Biosystems, CA, USA). The amplification conditions were as follows: 95°C for five minutes; followed by 40 cycles of 95°C for 15 seconds, 60°C for 30 seconds, and 72°C for 30 seconds; and a final five minute 72°C extension. All the primers were designed and synthesized by Takara Biomedical Technology (Beijing) Co., Ltd. and they are shown in Table 1.

### Statistical analysis

SPSS version 17.0 (SPSS Inc., Chicago, IL, USA) was used for all statistical analyses, and data were expressed as the mean ±SEM. The t-test was used to compare both two groups, and p values <0.05 were considered statistically significant.

### Results

#### Identification of DEGs

After data preprocessing, a total of 392 DEGs were identified in DEGs-1. Among these genes, 239 were upregulated and 153 were downregulated (Figure 1A). At the same time, the DEGs-2 set contained 332 DEGs (188 upregulated and 144 downregulated), and the top 20 DEGs of are shown in Figure 1B. The hierarchical clustering of DEGs in CRC samples, normal samples, miR-34a transformed CRC samples, and non-transfected CRC samples are shown in Figure 1C and 1D, respectively. The top 20 DEGs of DEGs-1 and DEGs-2 were separately listed in Table 2A and Table 2B.

#### Function enrichment of DEGs

There were 75 enriched GO terms of DEGs-1 identified, including 49 biological processes (BP), 19 cellular component (CC) and seven molecular function (MF) terms. Then, one KEGG pathway related to basal transcription factors was obtained. In DEGs-2, 243 GO terms and eight KEGG pathways were obtained. The top 10 GO terms for DEGs-1 and DEGs-2 are shown in Figure 2. Table 3 shows the enriched KEGG pathways of DEGs-1 and DEGs-2.

### Analysis of regulatory network of miR-34a in CRC

There were 1,668 target genes of miR-34a obtained via the miRWalk database, and 21 overlaps with DEGs-2 were identified, among which 11 genes (NNT4, KLF4, CEBPG, SEMA4B, PPARG, SERPINE1, C11orf38, FGD6, THBD, RTN1, and RRAGD) were downregulated. At the same time, six overlapped genes (KNTC1, CDC45, HAT1, DLAG5, KIAA0101, and FAM64A) were identified in the DEGs-1 and DEGs-2. Figure 3 shows the regulatory network between miR-34a and these overlapped genes.

#### Verification of important genes

Several genes such as SERPINE1 (downregulated), KLF4 (downregulated), SEMA4B (downregulated), PPARG (downregulated), CDC45 (upregulated) and KIAA0101 (upregulated) might be the targets of miR-34a. Results of RT-PCR are shown in Figure 4, which verified the related upregulated and downregulation of the aforementioned DEGs in CRC cells transformed with miR-34a compared with those non-transfected cells.

### Discussion

Despite advances in technologies of detection and therapies, CRC is still an uncontrollable disease. The lifetime risk of developing CRC is about 5.1% [26]. In the United States, CRC accounts for approximately 10% of cancer cases and cancer-related deaths [4]. MiR-34a was found to be a tumor suppressor gene.
in various cancers. In our study, DEGs, GO terms, and KEGG pathways in CRC samples and miR-34a transformed CRC samples were identified. Through the construction of the regulatory network of miR-34a, key genes in the progression of CRC were obtained, which might help make better understanding of the molecular mechanism and provide important reference for the diagnosis and therapy of CRC.

GO functional enrichment analysis indicated that both DEGs-1 and DEGs-2 were enriched in cell differentiation and cell cycle related biological processes. These processes were reported to be closely related to cancers. The majority of human cells were not cycling, while the minority of cells were cycling and were mainly located in self-renewing tissues. Deregulation of the cell cycle under laid the aberrant cell proliferation that characterized cancer [27], and tumor cells could not receive sufficient mitogenic signaling to drive them through the cycle and the division. At the same time, cell cycle phases could also be prognostic markers and therapy targets in various of cancers [28,29]. DEGs-2 has also been shown to be related to the biosynthesis of proteins and some other substances. Studies have shown that the rate of protein synthesis was a key factor

Figure 1. The DEGs change trend (A, B) and hierarchical cluster analysis of samples (C, D) in the data set of GSE4988 (A, C) and GSE7754 (B, D). Colors in the heat map represent for different gene expression level. Based on DEGs-1, CRC samples and normal samples were classified to different clusters (C). Similarly, miR-34a transformed CRC samples and non-transfected CRC samples were classified to different clusters according to DEGs-2 (D).
The top 20 DEGs of CRC samples compared to normal samples.

| Gene name | P value   | LogFC  |
|-----------|-----------|--------|
| DDx46     | 4.49×10⁻⁵ | 2.04625|
| ARM1C     | 0.000199  | 3.348333|
| HNRNPH3   | 0.000284  | 1.926875|
| FAM46A    | 0.000446  | -2.04625|
| EPAS1     | 0.000526  | 2.98625|
| HNRNPL    | 0.000576  | 2.830833|
| CRAT      | 0.000714  | 3.89375|
| TNKS      | 0.000876  | -1.9825|
| GPR98     | 0.0009    | -2.80917|
| CCNT1     | 0.00093   | -2.43167|
| CB1       | 0.001027  | 1.990417|
| DGKZ      | 0.001154  | 1.722083|
| CD83      | 0.001174  | 2.910417|
| BAD       | 0.001219  | -1.9975|
| IPO11     | 0.001407  | 1.87375|
| TRIP10    | 0.001578  | -1.94583|
| SMN2      | 0.001641  | 1.835|
| RNFI39    | 0.001666  | 2.78625|
| EIF3A     | 0.001693  | 3.027083|
| MCL1      | 0.001921  | 2.909583|

DEGs – differentially expressed genes; FC – fold change; CRC – colorectal cancer.

The top 20 DEGs of miR-34a transformed CRC samples compared to miR-34a non-transfected CRC samples.

| Gene name | P value   | LogFC  |
|-----------|-----------|--------|
| RRM2      | 2.82×10⁻⁸ | 3.34178|
| NCAPG     | 6.47×10⁻⁹ | 2.849829|
| PBK       | 7.61×10⁻⁹ | 2.939017|
| ANLN      | 7.74×10⁻⁹ | 2.775134|
| NMU       | 9.14×10⁻⁹ | 2.73189|
| TEMEL15850| 9.14×10⁻⁹ | -2.68421|
| CCNB1     | 9.31×10⁻⁹ | 2.897966|
| KIF11     | 9.33×10⁻⁹ | 2.829634|
| ANKRD29   | 9.56×10⁻⁹ | -2.66631|
| LURAP1L   | 1.35×10⁻⁸ | -2.56479|
| CDC20     | 1.50×10⁻⁸ | 2.758017|
| ZWINT     | 1.52×10⁻⁸ | 2.640488|
| MAD2L1    | 2.08×10⁻⁸ | 2.402782|
| SHCBP1    | 2.32×10⁻⁸ | 2.956084|
| PRC1      | 2.59×10⁻⁸ | 2.241771|
| MND1      | 2.87×10⁻⁸ | 2.368264|
| KIF14     | 3.04×10⁻⁸ | 2.958099|
| STK39     | 3.52×10⁻⁸ | 2.22861|
| MKI67     | 3.73×10⁻⁸ | 2.522868|

DEGs – differentially expressed genes; FC – fold change; CRC – colorectal cancer.

Protein synthesis was a significant process of metabolism and was closely related to many biological processes, such as cell cycle and cell differentiation. Studies showed that inhibition of mitochondrial protein synthesis lead to the lack of oxidative ATP generating capacity, which resulted in proliferation arrest of normal and malignant cells [31]. KEGG enrichment analysis showed that these DEGs were mainly related to cell cycle related pathway, pyrimidine metabolism related pathway, oocyte meiosis related pathway and DNA replication related pathway. Overall, they were all associated with cell growth, cell invasion, cell proliferation, and cell cycle, all of which play critical roles in the process of tumorigenesis [32]. Research has shown that miR-34a suppress tumors and inhibits recurrence of CRC through inhibiting cell growth, migration, and invasion, inducing cell apoptosis and cell cycle arrest [22,33]. Several target genes involved in these processes have been reported to be regulated by miR-34a so as to affect the pathogenetic process of CRC, such as SIRT1 and NOTCH1 [16,34]. The GO terms and KEGG pathways obtained in our study were in accordance with these other studies and indicated the function of miR-34a in regulating the progression of CRC.

There were 21 overlapped genes identified in DEGs-2 and the targets of miR-34a, among which 11 genes were downregulated in miR-34a transformed CRC samples, including NTN4, KLF4, CEBSG, SEMA4B, PPARG, SERPINE1, C11orf38, FGD6, THBD, RTN1, and RAGD. MiR-34a was identified to be a tumor suppressor in many types of solid tumor. It was reported that microRNAs usually act as a silencer of gene expression by binding to the 30 untranslated regions (30 UTRs) of target mRNAs, inhibiting their translation or marking them for degradation [17], so that the downregulated genes might be more reliable targets of miR-34a in the progression of CRC. Among...
Table 3. The enriched KEGG pathway for DEGs of CRC samples compared to normal samples, as well as DEGs of miR-34a transformed CRC samples compared to miR-34a non-transfected CRC samples.

| Category                      | Pathway name                        | Gene number | P value      |
|-------------------------------|-------------------------------------|-------------|--------------|
| KEGG pathway for DEGs-1      | Basal transcription factors         | 11          | 0.025411     |
| KEGG pathways for DEGs-2     |                                     |             |              |
| KEGG pathway                  | Cell cycle                          | 23          | 1.39 × 10^{-14} |
| KEGG pathway                  | DNA replication                      | 13          | 6.43 × 10^{-12} |
| KEGG pathway                  | Oocyte meiosis                       | 12          | 2.69 × 10^{-05} |
| KEGG pathway                  | Pyrimidine metabolism                | 10          | 2.28 × 10^{-04} |
| KEGG pathway                  | Progesterone-mediated oocyte maturation | 8         | 0.002809     |
| KEGG pathway                  | p53 signaling pathway                | 7           | 0.003764     |
| KEGG pathway                  | pathways in cancer                   | 15          | 0.013212     |
| KEGG pathway                  | Base excision repair                 | 4           | 0.041781     |

DEGs-1 – differentially expressed genes in colorectal samples compared to normal samples; DEGs-2 – differentially expressed genes in miR-34a transformed colorectal samples compared to miR-34a non-transfected colorectal samples. KEGG – Kyoto Encyclopedia of Genes and Genomes.
these genes, NTN4, KLF4, SERPINE1, and SEMA4B played critical roles in regulating cell growth, migration, and invasion in various of cancers, including CRC [9,35–37]. Studies showed that the expression of SERPINE1 was increased in CRC and was related to tumor invasiveness and aggressiveness [36]. KLF4 was an important factor in regulating cell cycle [38], which played important roles in the progression of CRC. Other genes such as CEBPG and PPARG were also reported to be related to the progression of cancers [39,40]. Many of these genes, such as KLF4 and PPARG, were reported to be regulatory targets of miR-34a [41,42]. There were studies that showed that miR-34a regulated apoptosis in liver cells by targeting the KLF4 gene [41]. These results indicated the critical roles of miR-34a in CRC, and indicated that miR-34a could affect the progression of CRC by regulating the expressions of genes related to cell migration, invasion, and other tumor related functions.

At the same time, six genes (KNTC1, CDC45, HAT1, DLGAP5, KIAA0101, and FAM64A), were identified to be overlaps between DEGs-1 and DEGs-2, indicating their potential functions in the progression of CRC and their relationship with miR-34a. These genes were also reported to be biomarkers or to be closely linked to cell migration, invasion, or DNA replication in various of cancers [43–46]. KIAA0101 was a p15PAF (proliferating cell nuclear antigen (PCNA)-associated factor) to bind with PCNA. Studies have shown that KIAA0101 was overexpressed in pancreatic cancer cells, and knocking down of KIAA0101 by small interfering RNA in pancreatic cancer cells caused drastic attenuation of cell proliferation, well exogenous overexpression of KIAA0101 enhanced cancer cell growth [45]. CDC45 plays a critical role in DNA replication. Studies showed that the expression of CDC45 was tightly associated with proliferating cell populations, and CDC45 seemed to be a promising candidate for a novel proliferation marker in cancer cell biology [46]. In our study, these overlapped genes were differently expressed in DEGs-1, DEGs-2, and the targets of miR-34a; thus we surmised that miR-34a might affect the progression of CRC by regulating the expression of these tumor related genes directly or indirectly.

**Conclusions**

In summary, pathways related to cell cycle, DNA replication, oocyte meiosis, and pyrimidine metabolism might be associated with CRC. MiR-34a plays an important role in regulating the progression of CRC and may provide an important reference for the diagnosis and treatment of CRC. Several genes such as SERPINE1, KLF4, SEMA4B, PPARG, CDC45, and KIAA0101 might be the targets of miR-34a and the potential therapeutic...
targets of CRC. Our study helps provide a better understanding of miR-34a in CRC and may provide important reference for the diagnosis and treatment of CRC. However, further experiments are still needed to confirm the results and to explore the specific regulation mechanism between miR-34a and these targets.

References:

1. Jemal A, Bray F, Center MM et al: Global cancer statistics, 2012. Cancer J Clin, 2011; 61: 69–90
2. Greenlee RT, Murray T, Bolden S, Wingo PA: Cancer statistics, 2000. Cancer J Clin, 2000; 50: 16–42
3. Pasetto LM, Monfardini S: Colorectal cancer screening in elderly patients: When should be more useful? Cancer Treat Rev, 2007; 33: 528–32.
4. Mclearney NL, Meyerhardt JA, Green E et al: Impact of age on the efficacy of newer adjuvant therapies in patients with stage II/III colon cancer: findings from the ACCENT database. J Clin Oncol, 2013; 31: 2600–6
5. Obrando DI, Gordon PH: Incidence and patterns of recurrence following curative resection for colorectal carcinoma. Dis Colon Rectum, 1997; 40: 15–24
6. André T, Boni C, Mounehjiboudiaf L et al: Oxaliplatin, fluorouracil, and leucovorin as adjuvant treatment for colon cancer. New Engl J Med, 2004; 350: 2343–51
7. Marisa I, De RA, Duval A et al: Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. PLoS Med, 2013; 10: e1001453
8. Carthew RW: Gene regulation by microRNAs. Curr Opin Genet Dev, 2006; 16: 203–8
9. Zhang J, Zheng F, Yu G et al: miR-196a targets netrin 4 and regulates cell proliferation and migration of cervical cancer cells. Biochem Biophys Res Commun, 2013; 440: 582–88
10. Schickel R, Boyerinas B, Park SM, Peter ME: MicroRNAs: Key players in the immune system, differentiation, tumorigenesis and cell death. Oncogene, 2008; 27: 5959–74
11. Ma Y, Li W, Wang H: Roles of miRNA in the initiation and development of colorectal carcinoma. Curr Pharm Des, 2012; 19: 1253–61
12. Rokavec M, MG Ö, Li H et al: Corrigendum. IL-6R/STAT3/miR-34a feedback loop promotes EMT-mediated colorectal cancer invasion and metastasis. J Clin Invest, 2014; 124: 1853–67
13. Lujambio A, Lowe SW: The microcosmos of cancer. Nature, 2012; 482: 347–55
14. Nugent M, Miller N, Kerin Mi: Circulating miR-34a levels are reduced in colorectal cancer. J Surg Oncol, 2012; 106: 947–52
15. Hermeking H: The miR-34 family in cancer and apoptosis. Cell Death & Differentiation, 2009; 17: 193–99
16. Bu P, Chen KY, Chen JH et al: A microRNA miR-34a-regulated bimodal switch targets notch in colon cancer stem cells. Cell Stem Cell, 2013; 12: 602–15
17. He L, He X, Lim PL et al: A microRNA component of the p53 tumour suppressor network. Nature, 2007; 447: 1130–34
18. Lapointe LC, Dunne R, Brown GS et al: Map of differential transcript expression and these targets. For their enthusiastic participation in this study.

Conflict of interests
None.

Acknowledgements
We would like to thank all the members of our research group for their enthusiastic participation in this study.

© Med Sci Monit, 2017; 23: 5735-5743
44. Loo LW, Cheng J, Tiirikainen M et al: cis-Expression QTL analysis of established colorectal cancer risk variants in colon tumors and adjacent normal tissue. PLoS One, 2012; 7: e30477

45. Hosokawa M, Takehara A, Matsuda K et al: Oncogenic role of KIAA0101 interacting with proliferating cell nuclear antigen in pancreatic cancer. Cancer Res, 2007; 67: 2568–76

46. Pollok S, Bauerschmidt C, Sänger J et al: Human Cdc45 is a proliferation-associated antigen. FEBS J, 2007; 274: 3669–84