

SCIENTIFIC REPORTS



OPEN

Colorectal cancer characterization and therapeutic target prediction based on microRNA expression profile

Received: 18 September 2015

Accepted: 08 January 2016

Published: 08 February 2016

Peng Xu, Yanliang Zhu, Bo Sun & Zhongdang Xiao

Colorectal cancer (CRC) is one of the most commonly diagnosed cancers and a major cause of cancer death. However, the molecular mechanisms underlying CRC initiation, growth and metastasis are poorly understood. In this study, based on our previous work for comprehensively analyzing miRNA sequencing data, we examined a series of colorectal cancer microRNAs expression profiles data. Results show that all these CRC samples share the same four pathways including TGF-beta signaling pathway, which is important in colorectal carcinogenesis. Twenty-one microRNAs that evolved in the four overlapped pathways were then discovered. Further analysis selected miR-21 as an important regulator for CRC through TGF-beta pathways. This study develops methods for discovering tumor specific miRNA cluster as biomarker and for screening new cancer therapy targets based on miRNA sequencing.

Colorectal cancer (CRC) is a major cause of cancer-associated morbidity and mortality all over the world, with an estimated 1.4 million cases and 693,900 deaths occurring in 2012 according to the *Global cancer statistics*^{1,2}. Invasion and metastasis are the major cause of this deadly cancer type. Although recent improvements in diagnostic techniques and clinical management have increased the early detection rate of CRC and decreased the mortality rate, more than one third of patients with CRC will ultimately develop into metastatic disease³. Until now, the pathogenesis of CRC is one of the most characterized carcinogenesis process, while new molecules that play a role in this process are still being discovered.

MicroRNAs (miRNAs) are small (approximately 20 ~ 22 nucleotides), noncoding RNAs that regulate gene expression by the inhibition of the translation and/or decreasing of the stability of their mRNA targets. It has been demonstrated that miRNAs play a significant role in tumorigenesis by down-regulating tumor suppressor genes or oncogenes^{4,5}. MiRNAs were first discovered in 1993, after that approximately two thousand human miRNAs have been identified, which have been proved to play important roles in various biological processes, including cell proliferation, cell cycle, apoptosis and differentiation⁶. An increasing number of miRNAs have been observed in various types of cancer and they were considered involving in modulating cancer cell behaviors^{7,8}. Indeed, dysregulating of miRNAs were proved to be a key process in the pathogenesis of many cancers and may occur at any point, from initiation to metastasis stage^{9,10}. Many studies have indicated that miRNAs were deregulated in various types of human malignancies such as colon cancer, breast cancer, lung cancer, pancreatic cancer, and chronic lymphocytic leukemia^{11,12}. Circulating levels of miRNAs are also quite stable and reproducible in the body fluid of many cancer patient. Given the critical role of miRNAs in the development of tumors and their global regulation, it shows great promise to character or screen therapy targets for CRC¹³.

In complex disease analysis for gaining potential delicate biological insights, pathway-based approaches and sub-pathway-based approaches have become the first choice¹⁴⁻¹⁶. In this study, based on the mathematical model as presented in our previous work¹⁷, we comprehensively analyze four CRC samples aiming to try to characterize CRC and discover therapeutic targets based on miRNAs profile.

Materials and Methods

Data source. The miRNA microarrays that we used in this study were deposited in the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus (GEO) database with accession code GSE35602,

State Key Laboratory of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, P. R. China. Correspondence and requests for materials should be addressed to B.S. (email: sunbo@seu.edu.cn) or Z.X. (email: zdxiao@seu.edu.cn)

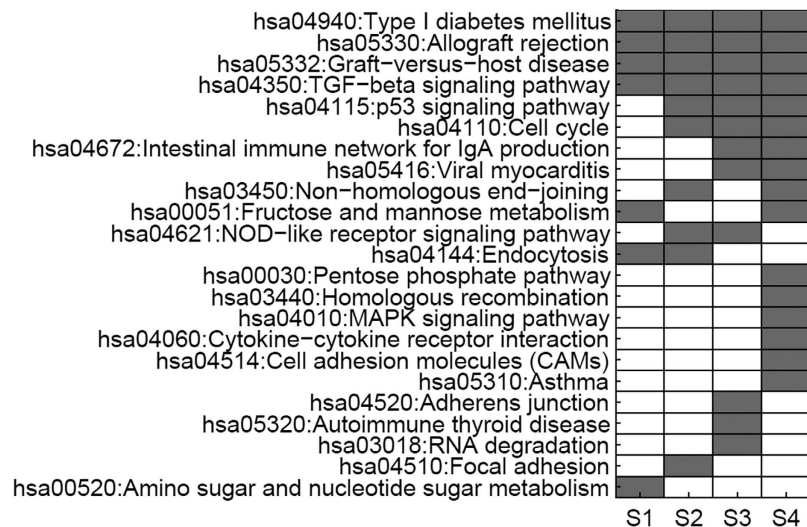


Figure 1. Significantly enriched KEGG pathways of submitted genes that significantly regulated by miRNAs in four CRC samples (S1, S2, S3 and S4). Gray filled block indicates genes are significantly enriched in this pathway in corresponding sample while white filled block indicates genes are not enriched in this pathway.

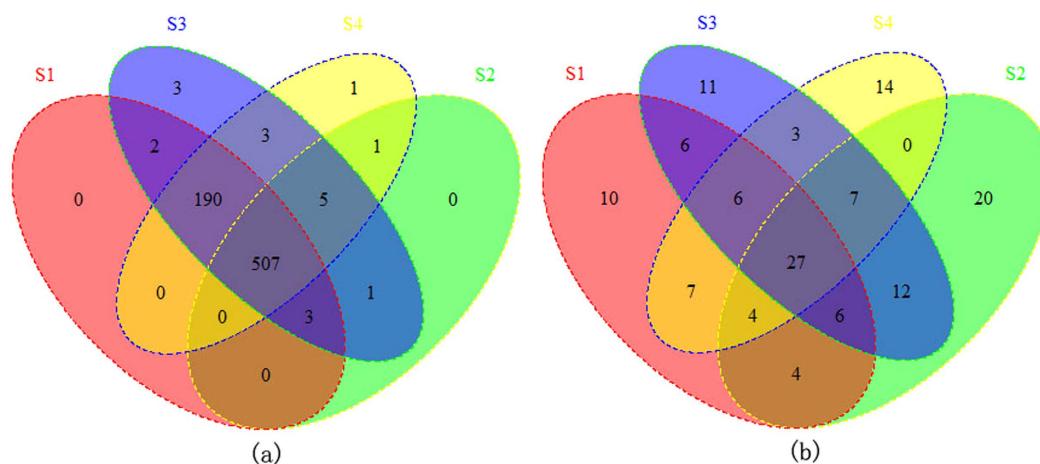


Figure 2. Venn Diagram of miRNAs correlated to the pathways respectively in four samples. All miRNAs that will target to the significantly repressed mRNAs occurring in enriched pathways respectively in four samples (a); miRNAs were filtered out after doing differential analysis on the repressed effect for miRNA-mRNA interactions (b).

which including four cases of CRC samples(GSM871528, GSM871531, GSM871533, GSM871535) and four cases of adjacent normal colorectal samples(GSM871539, GSM871541, GSM871543, GSM871545), as reported in¹⁸. In this study, total RNAs from epithelial tissue of cancer and normal samples were analyzed by miRNA microarray. Total RNAs were extracted from tissue using the miRNeasyMini Kit (Qiagen) according to the manufacturer's protocol. Concentrations and purities of the total RNAs were assessed with a spectrophotometer and RNA integrity was verified using an Agilent 2100 Bioanalyzer (Agilent Technologies). Colon adenocarcinoma (COAD) miRNASeq data samples were obtained from The Cancer Genome Atlas (TCGA).

Methods

miRNA expression profiles data normalization. Measures of RNA abundance are important for many areas of biology and often obtained from high throughput RNA sequencing methods such as Illumina sequence data. So in this study, when receive miRNA expression profiles data, to remove technical biases inherent in the sequencing approach we first normalize the expression profiles data. Here we employ TPM(Transcripts per million)¹⁹ method to do this assignment. The TPM of a miRNA *a* is:

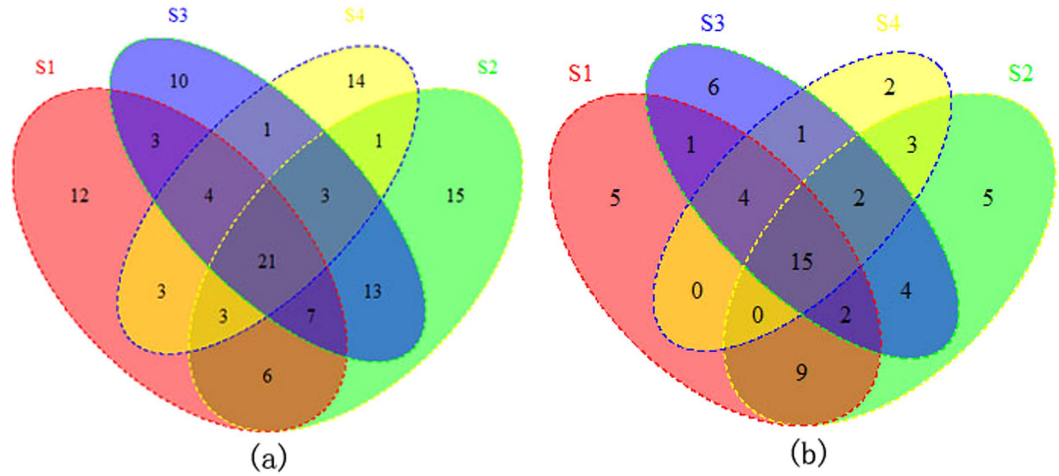


Figure 3. Venn Diagram of miRNAs (a) and mRNAs (b) correlated to the overlapped four pathways respectively in four samples.

$$TPM_a = \frac{[miRNA_a]}{[miRNA_a]} * 10^6 \tag{1}$$

$$[miRNA_{total}] = \sum_{a \in A} [miRNA_a] \tag{2}$$

where A stands for the set of all miRNAs determined in that experiment and a is an index for a miRNA. So both in eq.(1) and (2), the $miRNA_a$ represents transcripts number of miRNA a and $miRNA_{total}$ represents all miRNAs transcripts number in total in a sample.

A mathematical model to evaluate the comprehensive repression rate of specific mRNAs using total miRNA expression profiles. The method of evaluating the comprehensive repression rate of specific mRNAs using total miRNA expression profiles is as following¹⁷:

$$Re_{mRNA\alpha} = K * \frac{\sum_{i=a}^n [miRNA_a] * (RS_i - \alpha)^2}{\sum_{i=a}^n [miRNA_{total}] * \sum_{i=\alpha}^n RS_{a-i}} \tag{3}$$

Where $miRNA_a$ and $miRNA_{total}$ are the same meaning as refer to in above eq.(1) and (2). $RS_{i-\alpha}$ represents repression score of $miRNA_i$ on their target gene $mRNA_\alpha$, thereinto, i represents the indexes of miRNAs that share a specially target gene $mRNA_\alpha$. RS_{a-i} represents repression score of specific $miRNA_a$ on its target gene $miRNA_i$, here, the i represents the indexes of mRNAs that $miRNA_a$ will target on. Such repression score (RS) are extracted from miRanda²⁰.

Statistical analysis. Z-tests were performed to investigate differences in the repression effect of target mRNAs by miRNAs. p_0 was calculated using formula (4), which was used to investigate the null hypothesis.

$$p_0\alpha = \frac{Re\ mRNA\alpha_{control} + Re\ mRNA\alpha_{test}}{Re\ mRNA_{sum_control} + Re\ mRNA\alpha_{sum_test}} \tag{4}$$

Here, and $Re\ mRNA\alpha_{test}$ indicates the repression of $mRNA\alpha$ in the control and test groups, whereas $Re\ mRNA_{sum_control}$ and $Re\ mRNA\alpha_{sum_test}$ indicates the accumulated repression effect on mRNAs respectively in control and test groups.

Then, the Z-test was performed according to the following formula:

$$Z_\alpha = \frac{Re\ mRNA\alpha_{test}/Re\ mRNA_{sum_test} - Re\ mRNA\alpha_{control}/Re\ mRNA_{sum_control}}{\sqrt{p_0\alpha(1 - p_0\alpha)\left(\frac{1}{Re\ mRNA_{sum_control}} + \frac{1}{Re\ mRNA_{sum_test}}\right)}} \tag{5}$$

Method workflow.

- step 1: Normalization for miRNA expression profile;
- step 2: Compute the global repression effect of miRNA expression profile on their target genes;
- step 3: Do differential analysis for target genes based on the repression score;
- step 4: Do gene function enrichment analysis

miRNAs	up/down,References
hsa-miR-17*(hsa-miR-17-3p)	↑ ⁵³ .
	↑ ⁵⁴ .
	↑ ³⁵ .
hsa-miR-16-2*(hsa-miR-16-2-3p)	NULL
hsa-miR-30c-2*(hsa-miR-30c-2-3p)	NULL
hsa-miR-574-3p	↓ ⁵⁵ .
hsa-miR-592	NULL
hsa-miR-630	↑ ⁵⁶ .
hsa-miR-17	↑ ⁵⁷ .
	↓ ⁵⁸ .
hsa-miR-19b	↑ ⁵⁹ .
	↑ ⁶⁰ .
hsa-miR-21	↑ ³⁶ .
	↑ ³⁵ .
hsa-miR-22	↓ ⁶¹ .
hsa-miR-20a	↑ ⁵⁷ .
	↑ ⁵⁹ .
	↑ ⁶² .
	↑ ⁶³ .
hsa-miR-92a	↑ ⁶⁴ .
hsa-miR-29b	↑ ⁵⁹ .
hsa-miR-192	↓ ⁵⁷ .
	↓ ⁶⁵ .
hsa-miR-7	↑ ⁶⁶ .
hsa-miR-10b	↓ ⁵⁴ .
	↓ ⁵⁹ .
hsa-miR-215	↓ ⁵⁷ .
	↑ ³⁵ .
hsa-miR-224	↑ ⁶⁷ .
	↑ ⁵⁴ .
	↑ ⁵⁹ .
	↑ ³⁵ .
hsa-miR-150	↑ ⁶³ .
hsa-miR-194	↑ ³⁵ .
hsa-miR-20b	↓ ⁵⁴ .

Table 1. Reports of 21 overlapped miRNAs that significantly regulate the four pathways.

Results and Discussions

Four CRC tissue samples and normal colorectal epithelial tissue datasets were used in this study. The miRNAs expression profiles data for four normal colorectal epithelial were averaged among four samples as control group. Then four samples of CRC were respectively analyzed versus the control group. Currently, the methods for analysing miRNAs profiles are mainly according to the following steps, step 1: differential analysis for miRNAs expression profiles and select the differentially expressed miRNAs; step 2: predicting the targets genes of selected miRNAs; step 3: mapping the selected target genes to whole pathways to find significant enrichment pathways that related to the submitted genes list. However, it has been reported that an individual miRNA may target several mRNAs with different suppression level. The comprehensive repression of a given miRNA pool should be carefully calculated. Indeed, context scores for ranking the predicted targets of each miRNA have been provided by previous studies^{21–23}. To comprehensive analyze the global repression effect of miRNAs on their target genes, new method was bring forward in¹⁷.

Four groups of CRC and control samples were analyzed according to the method workflow mentioned in method section. Thereinto gene function enrichment analysis indicates four overlapped KEGG pathways (Type I diabetes mellitus, Allograft rejection, Graft-versus-host disease, TGF-beta signaling pathway) (see Fig. 1).

Then, we sort out all the significantly regulated mRNAs that participate in these pathways base on the reference provide by DAVID data base. Due to one miRNA can bind to many mRNAs and one mRNA can also be repressed by multiple miRNAs, hundreds of miRNAs will target to the significantly regulated mRNAs correlated to the overlapped pathways. Quite a few miRNAs (507) were overlapped between the four samples. The Venn Diagram of those mRNAs respectively in four samples of CRC versus control group were plotted in Fig. 2(a). Differential analysis was preformed after getting the repression effect of miRNAs on corresponding mRNAs in enriched pathways. Finally, significant differential miRNA-mRNA interactions were screened out. MiRNAs that

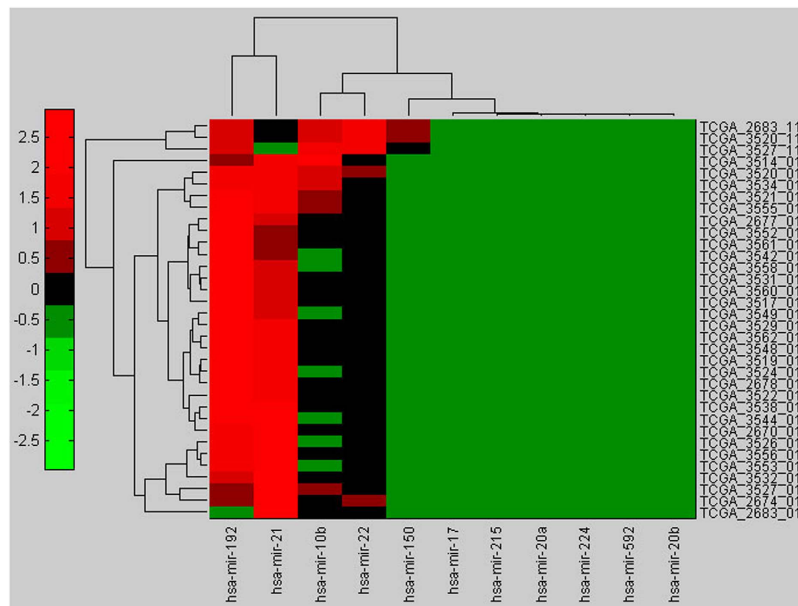


Figure 4. A panel of 11 miRNAs is sufficient to distinguish CRC patients from the healthy subjects. Identifiers for data samples were on the right side of the figure, thereinto, which end with ‘11’ are for healthy subjects and ‘01’ are for CRC samples.

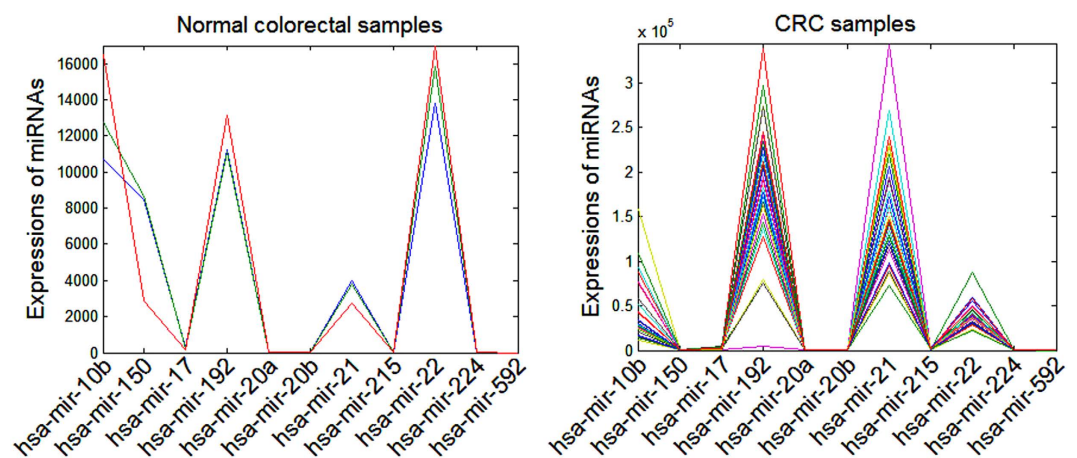


Figure 5. The expression profiles of 11 miRNAs respectively in normal colorectal and CRC samples.

occur in the significantly differential miRNA-mRNA interactions were thought to have relatively strong influence and significant differential repressed effect on their corresponding target gene. Results show that 27 miRNAs were overlapped between the four samples as described in Fig. 2(b).

To better understand the potential production mechanism of CRC, we narrow the searching space into four overlapped pathways. By doing differential analysis, results show that 21 miRNAs significantly regulate the mRNAs that correlated to four overlapped pathways as in Fig. 3(a). The mRNAs that were significantly regulated respectively in four samples were also obtained and statistical analysis was done as shown in Fig. 3(b).

For the 21 miRNAs screened out by above differential analysis, most of which also have been reported correlated to pathological mechanism of CRC in previous literatures list as in Table 1. To the best of our knowledge, 3 (3/21, 14.3%) miRNAs namely hsa-miR-16-2*, hsa-miR-30c-2* and hsa-miR-592 are without reports concerning with the CRC. Among the remaining 18 miRNAs, 16 (16/18, 88.9%) are respectively with an consistent change direction and 2 (2/18, 11.1%) with an inconsistent change direction among the published literatures (see Table 1).

As can be seen from Table 1, most of the 21 miRNAs were validated to be correlated to CRC, which suggests that these miRNAs may be potential biomarkers for CRC diagnosis. To verify this point of view, miRNASeq data samples for COAD stored in TCGA were adopted. Adenocarcinoma of the colon is the most common histopathological type of CRC, here we will uniformly called it as CRC. The miRNASeq data of CRC that we used here were in batch 28 including 31 CRC data samples and 3 matched normal samples. Firstly, we picked up miRNAs sequence data that belong to the 21 miRNAs, 12 miRNAs were selected, thereinto, miR-630 was found to be zero

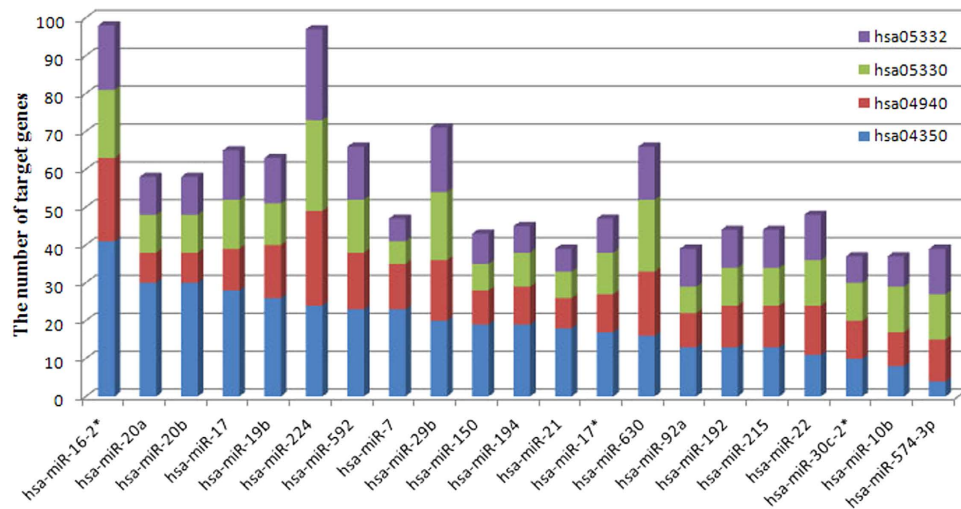


Figure 6. The number of target genes regulated by the 21 miRNAs respectively in the four overlapped pathways (Blue:hsa04350TGF-beta signaling pathway; Red:hsa04940Type I diabetes mellitus; Green: hsa05330Allograft rejection; Purple:hsa05332Graft-versus-host disease).

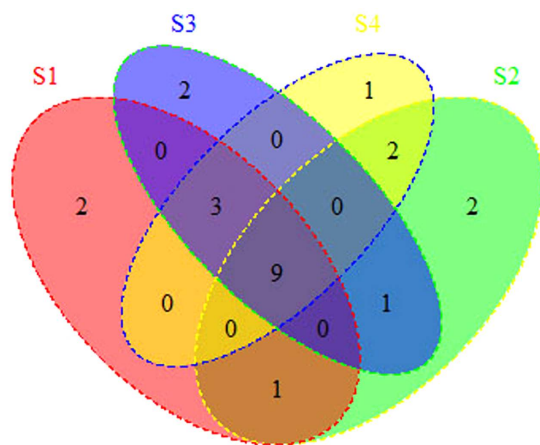


Figure 7. Venn Diagram significantly differential repressed mRNAs in hsa04350TGF-beta signaling pathway respectively in four CRC samples.

in all the 34 CRC miRNAseq samples. Ignored the miR-630, 11 miRNAs were remained in the data samples. Then, hierarchical clustering analysis in Matlab was adopted on the 34 CRC miRNAseq samples (see Fig. 4). Results show that a panel of 11 miRNAs is sufficient to distinguish CRC patients from the healthy subjects. The expression profiles of 11 miRNAs respectively in normal colorectal and CRC samples were described in Fig. 5. In normal colorectal samples, the expression of miR-21 is lower than miR-192 and miR-22 and expression of miR-22 is slightly higher than that of miR-192. In CRC samples, miR-21 and miR-192 are approximately equivalent in the expression and far more than that of miR-22. Compared with normal colorectal samples, both the expressions of miR-21 and miR-192 are upregulated in CRC samples, but the miR-22 is downregulated.

By KEGG pathway tool²⁴, we get the gene lists that evolved in different pathways. Thus, genes that presented in the overlapped four pathways respectively could be selected. Then from the miRNA-mRNA interaction reference given by database such as miRanda²⁰, we further get target genes regulated by the 21 miRNAs respectively in the four overlapped pathways (Fig. 6). Among those genes, as can be seen from the Fig. 6, most of those genes evolved in hsa04350TGF-beta signaling pathway. Furthermore, Type I diabetes mellitus, Allograft rejection and Graft-versus-host disease are signaling pathways that correlated to immune response. The reasons for immune response are various. Depending on these immune response pathways, it is difficult to discriminate whether the correlated dysregulated miRNAs are directly related to the CRC or are a secondary response in the histologic progression of the cancer. Bring together these factors, TGF-beta signaling pathway was regarded as the most important pathway that related to the CRC.

From Fig. 3(b), 15 significantly differential repressed mRNAs were overlapped between the four samples. Next, we paid close attention to the TGF-beta signaling pathway to see the distribution of 15 significantly differential

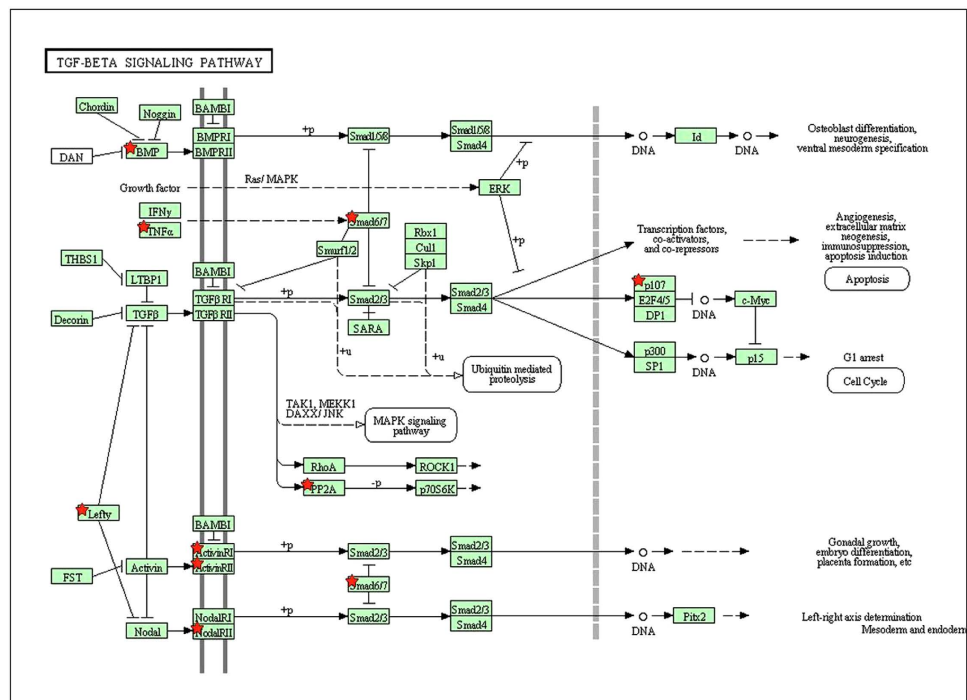


Figure 8. TGF-beta signaling pathway: the display of nine genes on the TGF-beta signaling pathway.

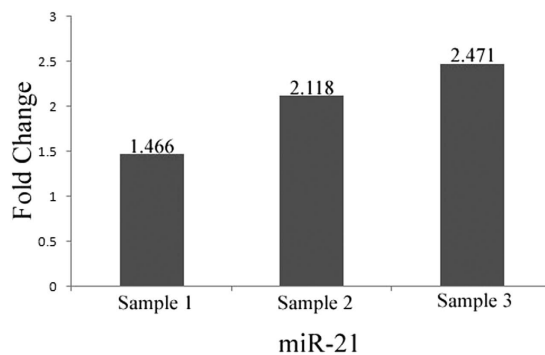


Figure 9. qPCR analysis of miR-21 expression level in three groups of CRC samples versus their corresponding control samples.

repressed mRNAs between the four samples (see Fig. 7). As can be seen from the diagram, there are nine genes evolved in all the four CRC samples. To study the potential biological function of the nine overlapped genes, we submit only the nine genes and see their distribution on the TGF-beta signaling pathway (see Fig. 8).

TGF-beta signaling pathway plays a unique and pivotal role in homeostasis, wound healing, carcinogenesis, fibrosis, angiogenesis and differentiation of the cell^{25,26}. TGF-beta and its signaling effectors influence cancer biological behaviors. The TGF-beta signaling pathway was proved to be a double-edged sword in cancer that plays as both a tumor suppressor and a cancer promoter²⁷. From the enriched genes in TGF-beta signaling pathway, we can see that genes Smad6/7, which are located in downstream of TNF-alpha, play as an inhibitory factor to Smad2/3. By the activated TGF- β RI receptor, Smad2/3 are phosphorylated at the carboxyl terminal serines and form heteromeric complexes with Smad4. Ultimately, Smad2/3/4 complex translocates into the nucleus and binds to specific regulatory elements on target genes^{28,29}. Smad 2/3/4 complex once in the nucleus, induces among others, a cyclin-dependent kinase (CDK) inhibitor p21, which will lead to growth arrest³⁰. Protein p21 is the product of waf/cip1 gene, an inhibitor of CDK, and is a powerful and reversible inhibitor of the propagation of cell cycle at G1 and G2 activated upon DNA damage. The protein p21 interacts with complexes of CDK2 and cyclin A and thereby inhibits the activity of CDK2, preventing progression of cell cycle³¹.

Indeed, we find that miR-21 has a significantly differential repressed effect on the Smad7 gene and miR-21 is up-regulated in all the four CRC samples respectively (Table in supplementary). In addition, Smad7 has been identified as a target gene of miR-21 namely in fibrotic lung diseases and diabetic nephropathy^{32,33}. The human miR-21 gene is located on chromosome 17q23-1 overlapping with the TMEM49 gene, a human homologue of rat

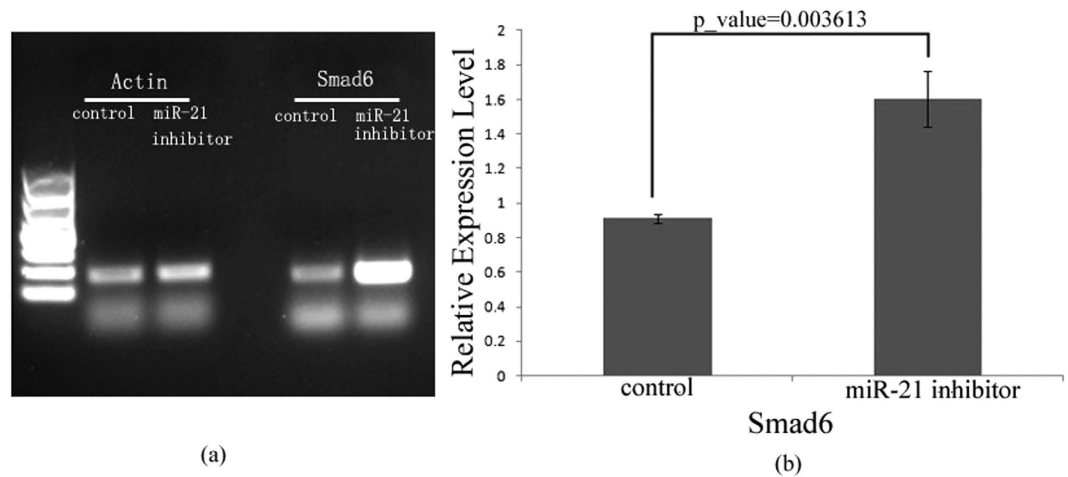


Figure 10. PCR analysis of expression level of Smad6 treated with and without inhibition of miR-21. PCR analysis were performed on Smad6, in which Actin was taken as internal reference (a). The same PCR analysis experiment were repeated 4 times, by using of software of imageJ, mean grey level of the PCR results were got and represented as mean \pm SD of four independent experiments (p_value = 0.003613) (b).

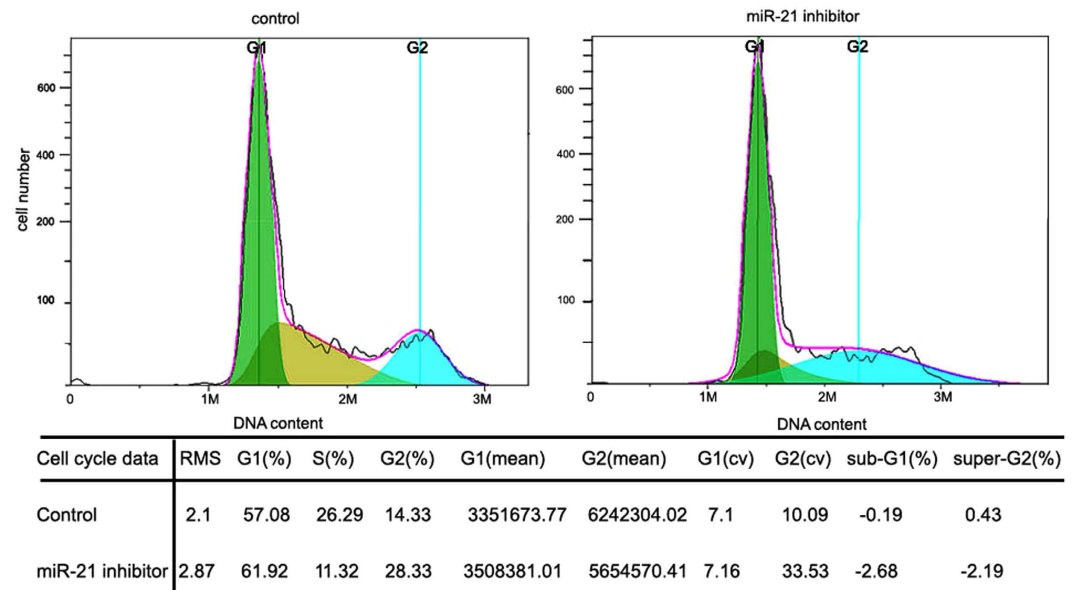


Figure 11. Effects of inhibition of miR-21 on cell cycle distribution in human CRC cell lines.

Vacuole Membrane Protein 1 (VMP1)³⁴. There have been several studies indicated that miR-21 expression was significantly higher in CRC compared with the adjacent normal tissues^{35–39}. Li *et al.* demonstrate that miR-21 and Smad7 are critical regulators of TGF- β 1 signaling during the induction of carcinoma associated fibroblasts(CAF) formation⁴⁰. They find that the overexpression of Smad7 or the depletion of miR-21 blocks TGF- β 1-induced CAF formation, whereas the overexpression of miR-21 or the depletion of Smad7 promotes CAF formation, even without TGF- β 1 stimulation. Carcinoma associated fibroblasts are activated fibroblasts and a key cellular component of tumor stroma. Through specific communication with cancer cells, CAFs not only directly promote tumor progression^{41,42} and metastasis^{43,44}, but are also involved in the initiation of cancer^{45–47}. Several studies have reported that the infection of pre-miR-21 lentiviral vector in CRC cell lines including HT-29, Colo206f, LIM 1863, SW480 and DLD1 cells can induce significant cell proliferation, enhance migratory and invasive properties, and decrease apoptosis. In contrast, knockdown of miR-21 could significantly decrease cell proliferation, suppress the invasion, migration and increase the apoptosis^{48–50}. Sprouty2 (Spry2) was considered as a tumor suppressor gene in cancer cells. The expression of Spry2 will inhibit the growth of HCT-116 cells, promote apoptosis and induce sensitization to 5-fluorouracil and metformin. But the knockdown of miR-21 in HCT-116 cells correlated

with the elevated expression of Spry2, and the proliferation rate of HCT-116 was down-regulated, indicating that function of Spry2 in CRC is regulated by miR-21⁵¹.

Validation on wet experiments. It inspired us that MiR-21 may be regarded as a target for inhibition the outgrowth of CRCs according the above data analysis. To verify this point, some wet experiments were carried out. Three CRC tissue samples and their corresponding paraneoplastic healthy tissue samples were adopted in this section. To dissect the expression difference of miR-21 between CRCs and healthy samples, qPCR analysis was done on the three groups of samples. Results show that the expression level of miR-21 are higher in all cancer samples than that in control groups with fold changes 1.466, 2.118 and 2.471 respectively (see Fig. 9). These results are consistent with the conclusion obtained from above studies and indicates miR-21 could be regarded as a target in CRC treatment.

As described in Fig. 8, differentially repressed Smad6/7 locate in an important position which is correlated to both the cell cycle and embryo differentiation. Smad7 has been validated to be the target gene of miR-21 in previously experiments^{32,33} and was regard as a pivot factor in cancer. Interestingly, though both Smad6 and Smad7 are the targets of miR21, there is lack of reports that miR-21 could regulate Smad6 in colon cancer cells, to the best of our knowledge. To examine the Smad6 expression level treated with and without inhibition of miR-21, PCR analysis were implemented. As our expected, results show that the expression level of Smad6 is becoming significantly higher when inhibit the level of miR-21 in colon cancer cells (see Fig. 10).

Finally, to examine the effect of miR-21 on colon cancer cell growth, miR-21 antisense was inducted into the cells, the cell cycle analysis was performed. As our expect, the growth of CRCs was significantly inhibited as shown in Fig. 11. These data suggest miR-21 inhibition in CRCs could effectively impair the growth of tumor cells and thus provide a therapeutic target for CRC treatment.

Real Time Quantitative Polymerase Chain Reaction (qRT-PCR). Total RNA of CRC tissue samples and their corresponding paraneoplastic healthy tissue samples were extracted using Trizol reagent according to the supplier's instruction (Invitrogen, USA) and quantified by spectrophotometry at 260 nm. Reverse transcription was performed for 60 min at 42 °C, followed by 5 min at 85 °C to inactivate the reverse transcriptase. MicroRNA expression level was examined in all samples. The approach of the experiment followed standard protocols, qRT-PCR was performed on the Applied Biosystems 7500 Sequence Detection System (BD, Michigan, USA). The reactions were amplified over 40 cycles of 95 °C for 15s and 60 °C for 1 min. Afterwards, a thermal denaturation protocol was performed to determine the number of products present in each reaction. The reactions were typically run in triplicate. The cycle number at which the reaction crossed an arbitrarily placed threshold (Ct) was determined for each gene. Target gene expression was normalized to the expression of U6 in each sample. Data was analyzed using the $2^{-\Delta\Delta Ct}$ method⁵².

Conclusions

Due to the stability of miRNAs during analysis and the development of sequencing methods, miRNA expression profiling was identified as a reliable method for investigation of mRNAs, proteins and pathways involved in particular biological processes. In this study, we comprehensively analyze a series of CRC miRNAs expression profiles data, 21 miRNAs that have significantly differential repressed effect on target genes were found. Gene function analysis show that TGF-beta signaling pathway may involved in colorectal carcinogenesis and miR-21 was found to be important in this process. Inhibition miR-21 could hinder the growth of CRC growth. Thus, through miRNA sequencing analyzing, we could characterize the CRC with a panel of 21 miRNAs, as well as select potential target for CRC treatment.

References

1. Edwards, B. K. *et al.* Annual Report to the Nation on the status of cancer, 1975-2010, featuring prevalence of comorbidity and impact on survival among persons with lung, colorectal, breast, or prostate cancer. *Cancer* **120**, 1290–1314, doi: 10.1002/cncr.28509 (2014).
2. Torre, L. A. *et al.* Global cancer statistics, 2012. *CA: a cancer journal for clinicians* **65**, 87–108, doi: 10.3322/caac.21262 (2015).
3. Siegel, R. *et al.* Cancer treatment and survivorship statistics, 2012. *CA: a cancer journal for clinicians* **62**, 220–241, doi: 10.3322/caac.21149 (2012).
4. Wienholds, E. & Plasterk, R. H. MicroRNA function in animal development. *FEBS Lett* **579**, 5911–5922, doi: 10.1016/j.febslet.2005.07.070 (2005).
5. Esteller, M. Non-coding RNAs in human disease. *Nat Rev Genet* **12**, 861–874, doi: 10.1038/nrg3074 (2011).
6. Kanellopoulou, C. & Monticelli, S. A role for microRNAs in the development of the immune system and in the pathogenesis of cancer. *Seminars in cancer biology* **18**, 79–88, doi: 10.1016/j.semcan.2008.01.002 (2008).
7. Lovat, F., Valeri, N. & Croce, C. M. MicroRNAs in the pathogenesis of cancer. *Seminars in oncology* **38**, 724–733, doi: 10.1053/j.seminoncol.2011.08.006 (2011).
8. Gotte, M. MicroRNAs in breast cancer pathogenesis. *Minerva ginecologica* **62**, 559–571 (2010).
9. Dela Cruz, F. & Matushansky, I. MicroRNAs in chromosomal translocation-associated solid tumors: learning from sarcomas. *Discovery medicine* **12**, 307–317 (2011).
10. Bartel, D. P. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**, 281–297 (2004).
11. Iorio, M. V. & Croce, C. M. MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. *EMBO molecular medicine* **4**, 143–159, doi: 10.1002/emmm.201100209 (2012).
12. An, J., Zhu, X., Wang, H. & Jin, X. A dynamic interplay between alternative polyadenylation and microRNA regulation: implications for cancer (Review). *Int J Oncol* **43**, 995–1001, doi: 10.3892/ijo.2013.2047 (2013).
13. Zeng, W. *et al.* Predictive power of circulating miRNAs in detecting colorectal cancer. *Tumour biology: the journal of the International Society for Oncodevelopmental Biology and Medicine* **36**, 2559–2567, doi: 10.1007/s13277-014-2872-2 (2015).
14. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 15545–15550, doi: 10.1073/pnas.0506580102 (2005).

15. Li, X., Shen, L., Shang, X. & Liu, W. Subpathway Analysis based on Signaling-Pathway Impact Analysis of Signaling Pathway. *PLoS one* **10**, e0132813, doi: 10.1371/journal.pone.0132813 (2015).
16. Vrahatis, A. G., Dimitrakopoulou, K., Balomenos, P., Tsakalidis, A. K. & Bezirianos, A. CHRONOS: A time-varying method for microRNA-mediated sub-pathway enrichment analysis. *Bioinformatics*, doi: 10.1093/bioinformatics/btv673 (2015).
17. Sun, B., Yang, F., Hu, F.-H., Huang, N.-P. & Xiao, Z.-D. Comprehensive annotation of microRNA expression profiles. *BMC genetics* **14**, 120 (2013).
18. Nishida, N. *et al.* Microarray Analysis of Colorectal Cancer Stromal Tissue Reveals Upregulation of Two Oncogenic miRNA Clusters. *Clin. Cancer Res.* **18**, 3054–3070, doi: 10.1158/1078-0432.ccr-11-1078 (2012).
19. Wagner, G. P., Kin, K. & Lynch, V. J. Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory in Biosciences* **131**, 281–285, doi: 10.1007/s12064-012-0162-3 (2012).
20. Betel, D., Koppal, A., Agius, P., Sander, C. & Leslie, C. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome biology* **11**, R90, doi: 10.1186/gb-2010-11-8-r90 (2010).
21. Ulitsky, I., Laurent, L. C. & Shamir, R. Towards computational prediction of microRNA function and activity. *Nucleic acids research* **38**, doi: 10.1093/nar/gkq570 (2010).
22. Friedman, R. C., Farh, K. K. H., Burge, C. B. & Bartel, D. P. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Research* **19**, 92–105 (2009).
23. Sahoo, S. & Albrecht, A. A. Ranking of microRNA target prediction scores by Pareto front analysis. *Comput Biol Chem* **34**, 284–292, doi: 10.1016/j.compbiolchem.2010.09.005 (2010).
24. Kanehisa, M. & Goto, S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic acids research* **28**, 27–30 (2000).
25. Derynck, R. & Feng, X. H. TGF-beta receptor signaling. *Biochimica et biophysica acta* **1333**, F105–150 (1997).
26. Massague, J. TGF-beta signal transduction. *Annual review of biochemistry* **67**, 753–791, doi: 10.1146/annurev.biochem.67.1.753 (1998).
27. Akhurst, R. J. & Derynck, R. TGF-beta signaling in cancer—a double-edged sword. *Trends in cell biology* **11**, S44–51 (2001).
28. Miyazono, K., ten Dijke, P. & Heldin, C. H. TGF-beta signaling by Smad proteins. *Advances in immunology* **75**, 115–157 (2000).
29. Derynck, R., Zhang, Y. & Feng, X. H. Smads: transcriptional activators of TGF-beta responses. *Cell* **95**, 737–740 (1998).
30. Lampropoulos, P. *et al.* TGF-beta signalling in colon carcinogenesis. *Cancer letters* **314**, 1–7, doi: 10.1016/j.canlet.2011.09.041 (2012).
31. Reynisdottir, I., Polyak, K., Iavarone, A. & Massague, J. Kip/Cip and Ink4 Cdk inhibitors cooperate to induce cell cycle arrest in response to TGF-beta. *Genes & development* **9**, 1831–1845 (1995).
32. Liu, G. *et al.* miR-21 mediates fibrogenic activation of pulmonary fibroblasts and lung fibrosis. *The Journal of experimental medicine* **207**, 1589–1597, doi: 10.1084/jem.20100035 (2010).
33. Zhong, X. *et al.* miR-21 is a key therapeutic target for renal injury in a mouse model of type 2 diabetes. *Diabetologia* **56**, 663–674, doi: 10.1007/s00125-012-2804-x (2013).
34. Zhao, J., Xu, G., Zhang, Y. & Zhao, G. Emerging role of microRNA-21 in colorectal cancer. *Cancer biomarkers : section A of Disease markers*, doi: 10.3233/CBM-150468 (2015).
35. Monzo, M. *et al.* Overlapping expression of microRNAs in human embryonic colon and colorectal cancer. *Cell Res* **18**, 823–833, doi: 10.1038/cr.2008.81 (2008).
36. Kulda, V. *et al.* Relevance of miR-21 and miR-143 expression in tissue samples of colorectal carcinoma and its liver metastases. *Cancer genetics and cytogenetics* **200**, 154–160, doi: 10.1016/j.cancergencyto.2010.04.015 (2010).
37. Schee, K., Boye, K., Abrahamson, T. W., Fodstad, O. & Flatmark, K. Clinical relevance of microRNA miR-21, miR-31, miR-92a, miR-101, miR-106a and miR-145 in colorectal cancer. *BMC Cancer* **12**, 505, doi: 10.1186/1471-2407-12-505 (2012).
38. Kanaan, Z. *et al.* Plasma miR-21: a potential diagnostic marker of colorectal cancer. *Annals of surgery* **256**, 544–551, doi: 10.1097/SLA.0b013e318265bd6f (2012).
39. Link, A. *et al.* Fecal MicroRNAs as novel biomarkers for colon cancer screening. *Cancer epidemiology, biomarkers & prevention : a publication of the American Association for Cancer Research, cosponsored by the American Society of Preventive Oncology* **19**, 1766–1774, doi: 10.1158/1055-9965.EPI-10-0027 (2010).
40. Li, Q. *et al.* MiR-21/Smad 7 signaling determines TGF-beta1-induced CAF formation. *Scientific reports* **3**, 2038, doi: 10.1038/srep02038 (2013).
41. Dimanche-Boitrel, M. T. *et al.* *In vivo* and *in vitro* invasiveness of a rat colon-cancer cell line maintaining E-cadherin expression: an enhancing role of tumor-associated myofibroblasts. *International journal of cancer. Journal international du cancer* **56**, 512–521 (1994).
42. Orimo, A. *et al.* Stromal fibroblasts present in invasive human breast carcinomas promote tumor growth and angiogenesis through elevated SDF-1/CXCL12 secretion. *Cell* **121**, 335–348, doi: 10.1016/j.cell.2005.02.034 (2005).
43. Grum-Schwensen, B. *et al.* Suppression of tumor development and metastasis formation in mice lacking the S100A4(mts1) gene. *Cancer research* **65**, 3772–3780, doi: 10.1158/0008-5472.CAN-04-4510 (2005).
44. Olaso, E. *et al.* Tumor-dependent activation of rodent hepatic stellate cells during experimental melanoma metastasis. *Hepatology* **26**, 634–642, doi: 10.1002/hep.510260315 (1997).
45. Olumi, A. F. *et al.* Carcinoma-associated fibroblasts direct tumor progression of initiated human prostatic epithelium. *Cancer research* **59**, 5002–5011 (1999).
46. Bhowmick, N. A. *et al.* TGF-beta signaling in fibroblasts modulates the oncogenic potential of adjacent epithelia. *Science* **303**, 848–851, doi: 10.1126/science.1090922 (2004).
47. Kuperwasser, C. *et al.* Reconstruction of functionally normal and malignant human breast tissues in mice. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 4966–4971, doi: 10.1073/pnas.0401064101 (2004).
48. Cottonham, C. L., Kaneko, S. & Xu, L. miR-21 and miR-31 converge on TIAM1 to regulate migration and invasion of colon carcinoma cells. *The Journal of biological chemistry* **285**, 35293–35302, doi: 10.1074/jbc.M110.160069 (2010).
49. Deng, J. *et al.* Targeting miR-21 enhances the sensitivity of human colon cancer HT-29 cells to chemoradiotherapy *in vitro*. *Biochemical and biophysical research communications* **443**, 789–795, doi: 10.1016/j.bbrc.2013.11.064 (2014).
50. Asangani, I. A. *et al.* MicroRNA-21 (miR-21) post-transcriptionally downregulates tumor suppressor Pdc4 and stimulates invasion, intravasation and metastasis in colorectal cancer. *Oncogene* **27**, 2128–2136, doi: 10.1038/sj.onc.1210856 (2008).
51. Tsao, C.-J. microRNA-21-mediated regulation of Sprouty2 protein expression enhances the cytotoxic effect of 5-fluorouracil and metformin in colon cancer cells. *International journal of molecular medicine*, doi: 10.3892/ijmm.2012.910 (2012).
52. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(T)(-Delta Delta C) method. *Methods* **25**, 402–408 (2001).
53. Ng, E. K. *et al.* Differential expression of microRNAs in plasma of patients with colorectal cancer: a potential marker for colorectal cancer screening. *Gut* **58**, 1375–1381, doi: 10.1136/gut.2008.167817 (2009).
54. Sarver, A. L. *et al.* Human colon cancer profiles show differential microRNA expression depending on mismatch repair status and are characteristic of undifferentiated proliferative states. *BMC Cancer* **9**, 401, doi: 10.1186/1471-2407-9-401 (2009).
55. Della Vittoria Scarpati, G. *et al.* Analysis of differential miRNA expression in primary tumor and stroma of colorectal cancer patients. *Biomed Res Int* **2014**, 840921, doi: 10.1155/2014/840921 (2014).
56. Chu, D. *et al.* MicroRNA-630 is a prognostic marker for patients with colorectal cancer. *Tumor Biol.* **35**, 9787–9792, doi: 10.1007/s13277-014-2223-3 (2014).

57. Earle, J. S. *et al.* Association of microRNA expression with microsatellite instability status in colorectal adenocarcinoma. *The Journal of molecular diagnostics : JMD* **12**, 433–440, doi: 10.2353/jmoldx.2010.090154 (2010).
58. Diosdado, B. *et al.* MiR-17-92 cluster is associated with 13q gain and c-myc expression during colorectal adenoma to adenocarcinoma progression. *British journal of cancer* **101**, 707–714, doi: 10.1038/sj.bjc.6605037 (2009).
59. Arndt, G. M. *et al.* Characterization of global microRNA expression reveals oncogenic potential of miR-145 in metastatic colorectal cancer. *BMC Cancer* **9**, 374, doi: 10.1186/1471-2407-9-374 (2009).
60. Kurokawa, K. *et al.* Role of miR-19b and its target mRNAs in 5-fluorouracil resistance in colon cancer cells. *Journal of gastroenterology* **47**, 883–895, doi: 10.1007/s00535-012-0547-6 (2012).
61. Tsuchiya, N. *et al.* Tumor suppressor miR-22 determines p53-dependent cellular fate through post-transcriptional regulation of p21. *Cancer research* **71**, 4628–4639, doi: 10.1158/0008-5472.CAN-10-2475 (2011).
62. Schepeler, T. *et al.* Diagnostic and prognostic microRNAs in stage II colon cancer. *Cancer research* **68**, 6416–6424, doi: 10.1158/0008-5472.CAN-07-6110 (2008).
63. Volinia, S. *et al.* A microRNA expression signature of human solid tumors defines cancer gene targets. *Proceedings of the National Academy of Sciences of the United States of America* **103**, 2257–2261, doi: 10.1073/pnas.0510565103 (2006).
64. Huang, Z. H. *et al.* Plasma microRNAs are promising novel biomarkers for early detection of colorectal cancer. *International Journal of Cancer* **127**, 118–126 (2010).
65. Chen, X. *et al.* Role of miR-143 targeting KRAS in colorectal tumorigenesis. *Oncogene* **28**, 1385–1392, doi: 10.1038/onc.2008.474 (2009).
66. Slattery, M. L. *et al.* MicroRNAs and colon and rectal cancer: differential expression by tumor location and subtype. *Genes, chromosomes & cancer* **50**, 196–206, doi: 10.1002/gcc.20844 (2011).
67. Wang, Y. X. *et al.* Initial study of microRNA expression profiles of colonic cancer without lymph node metastasis. *Journal of digestive diseases* **11**, 50–54, doi: 10.1111/j.1751-2980.2009.00413.x (2010).

Acknowledgements

This work was supported by grants from the Natural Science Foundation of Jiangsu Province (No. BK2012122), the National Natural Science Foundation of China (NSFC) (No. 61071047), the Major State Basic Research Development Program of China (973 Program) (No. 2013CB932902), and the Specialized Research Fund for the Doctoral Program of Higher Education of China (No. 20130092110030). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Author Contributions

Conceived and designed the experiments: B.S. and Z.D.X. Developed and implemented the analysis tools: P.X.; Performed the biological experiments: Y.L.Z. Analyzed the data: P.X., B.S. and Z.D.X. Wrote the paper: P.X., B.S. and Z.D.X. All authors read and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Xu, P. *et al.* Colorectal cancer characterization and therapeutic target prediction based on microRNA expression profile. *Sci. Rep.* **6**, 20616; doi: 10.1038/srep20616 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>