

# Systematic literature review and clinical validation of circulating microRNAs as diagnostic biomarkers for colorectal cancer

Cheng Pan<sup>1,2,\*</sup>, Xuebing Yan<sup>1,\*</sup>, Hao Li<sup>1,\*</sup>, Linsheng Huang<sup>1,\*</sup>, Mingming Yin<sup>1</sup>, Yongzhi Yang<sup>1</sup>, Renyuan Gao<sup>1</sup>, Leiming Hong<sup>1,3</sup>, Yanlei Ma<sup>1</sup>, Chenzhang Shi<sup>1</sup>, Huanlong Qin<sup>1</sup> and Peng Zhang<sup>1</sup>

<sup>1</sup>Department of General Surgery, Shanghai Tenth People's Hospital Affiliated to Tongji University No. 301, Shanghai 200072, China

<sup>2</sup>Medical Department, Soochow University, Jiangsu 215123, China

<sup>3</sup>Department of General Surgery, Weihai Municipal Hospital, Shandong 264200, China

\*These authors have contributed equally to this work

Correspondence to: Huanlong Qin, email: huanlong\_qin@live.cn

Peng Zhang, email: ppcon@126.com

**Keywords:** colorectal cancer, circulating microRNAs, diagnosis, biomarkers

**Received:** November 30, 2016

**Accepted:** June 20, 2017

**Published:** July 18, 2017

**Copyright:** Pan et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License 3.0 (CC BY 3.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

## ABSTRACT

**Because patients with colorectal cancer (CRC) are usually diagnosed at an advanced stage and current serum tumor markers have limited diagnostic efficacy, there is an urgent need to identify reliable diagnostic biomarkers. To better define the diagnostic potential of microRNAs (miRNAs) for CRC, we performed a comprehensive evaluation of reported circulating CRC miRNA markers. After a systematic literature review, we selected 30 candidate miRNAs and used quantitative real-time polymerase chain reaction to examine their expression in a training cohort of 120 plasma samples (CRC vs healthy controls (HC) = 60:60). Expression data was confirmed in a validation cohort of 160 plasma samples (CRC vs HC = 80:80). We ultimately identified 5 dysregulated circulating miRNAs (miR-15b, miR-17, miR-21, miR-26b, and miR-145), of which miR-21 and miR-26b proved to have the best diagnostic performance in the training and validation cohorts, respectively. Based on these results, we propose a novel blood-based diagnostic model, integrating 5 CRC-related miRNAs and serum carcinoembryonic antigen (CEA), which provides better diagnostic performance than the combined 5 miRNAs, CEA alone, or any single miRNA. We propose that the novel CRC diagnostic model presented here will be useful for overcoming the limitations faced by current non-invasive diagnostic strategies.**

## INTRODUCTION

Colorectal cancer (CRC) is a commonly diagnosed cancer worldwide, and its incidence is dramatically increasing in developing countries with growing aging population and westernized lifestyles. Although population-based screening has proved to effectively prevent CRC development when detected at early stage, approximately 60% of CRC patients are diagnosed at regional or distant stages, with a discouraging 5-year

survival rate ranging from 12.5% to 70.4% [1]. Moreover, when assessed alone, current tumor markers such as the carcinoembryonic antigen (CEA) and CA125 are frequently ineffective for early CRC detection, inevitably resulting in delayed diagnosis [2]. As numerous molecular biomarkers of CRC progression and prognosis have been recently proposed [3], there is substantial hope in combining them with conventional clinical parameters to more accurately diagnose CRC and guide treatment.

Several studies in the last few years have closely linked microRNAs (miRNAs) with the initiation and

development of various human malignancies [4]. Several studies from our research group have also assessed the biological roles and relevant mechanisms of diverse miRNAs in CRC tumorigenesis. For example, we identified miR-17 as an oncogenic miRNA that promotes CRC development by activating the Wnt/ $\beta$ -catenin pathway by targeting P130 [5]. We also found that miR-149 methylation contributes to CRC growth and invasion by targeting the transcription factor Sp1 [6]. Furthermore, using miRNA expression profiling, we identified miR-150 as a prognostic biomarker for chemotherapy response and defined its anti-cancer effects, exerted through c-myc downregulation [7, 8]. Lastly, we defined a novel oncogenic role for miR-21 in the malignant transformation of colitis-associated CRC, where it targets the tumor suppressor PDCD4, activating the pro-inflammatory NF- $\kappa$ B/STAT3 cascade [9].

Given the crucial roles played by miRNAs in CRC development, there is great potential in translating them into clinically actionable biomarkers for diagnosis and prognostication. Studies revealed that tumor-derived miRNAs can be released into the circulation by exosomes, microvesicles, or bound to RNA binding proteins and lipoproteins [10, 11]. In 2009 Ng et al. identified for the first time a plasma miRNA marker, miR-92, that distinguished CRC patients from healthy controls with a sensitivity of 89% and a specificity of 70% [12]. Following his work, numerous studies reported the diagnostic value of other circulating miRNAs, such as miR-21 and miR-221, in CRC patients [13-15]. However, these achievements have not been satisfactorily translated into clinical benefits largely due to insufficient retrospective validation on a highly standardized platform. Consequently, the clinical utility of combined miRNAs and their potential cooperation with traditional noninvasive diagnostic tumor markers (such as CEA) remain undetermined.

To address this issue, we performed a systematic literature review and selected candidate circulating miRNAs from relevant studies, including our own published work. Then, we employed a training cohort and a validation cohort to evaluate their diagnostic value. Finally, we constructed and validated a novel diagnostic model integrating multiple miRNAs and conventional tumor markers. These efforts not only provide a comprehensive evaluation of the diagnostic value of circulating miRNAs in CRC, but also strongly promote the clinical translation of a novel, non-invasive diagnostic approach.

## RESULTS

### Selection of candidate circulating miRNAs

We initially performed a systematic literature review and selected 82 miRNAs with diagnostic potential that were detected in CRC patients (Figure 1). To further narrow our candidate list, we first excluded miRNAs detected in tissues (n = 24) and feces (n = 10). We also

excluded miRNAs detected in whole blood (n = 5) because blood cells are a major contributor to circulating miRNA and may have an equivocal impact on analysis [16, 17]. We then excluded controversial miRNAs which were investigated by divergent methodologies (n = 9) and studies which enrolled fewer than 50 samples (n = 8). As a result, 26 miRNAs were preserved (Supplementary Table 1) [12, 14, 15, 18-39]. We added to this dataset 4 miRNAs (miR-17, miR-26b, miR-149, and miR-150) found to be involved in CRC development in our previous studies [5-7, 40]. Thus, a total of 30 miRNAs were finally selected for screening in the following training phase.

### Circulating miRNA screening and diagnostic model construction in the training phase

In the training phase, qRT-PCR was first performed to compare the expression of candidate circulating miRNAs between CRC patients (n = 60) and healthy controls (HC) (n = 60) (Table 1). As shown in Figure 2, the expression of 5 circulating miRNAs was significantly higher in CRC than in HC (miR-15b, p = 0.0005; miR-17, p = 0.0007; miR-21, p < 0.0001; miR-26b, p = 0.0001; miR-145, p = 0.0008). No significant differences were observed between CRC and HC in the other 25 circulating miRNAs [p > 0.0017 (0.05/30)].

Then, we depicted ROC curves to evaluate the performance of these miRNAs in discriminating CRC from HC. As shown in Table 2 and Figure 4A, miR-21 had the best discriminatory performance with an AUC of 0.708, while miR-145 had the worst performance with an AUC of 0.677.

Binary logistic regression analysis was conducted to combine the 5 dysregulated miRNAs with 5 commonly used tumor markers (CEA, CA19-9, CA-724, CA153, and CA125). As a result, two diagnostic models were constructed. One consists of only 5 miRNAs (miR-15b, miR-17, miR-21, miR-26b, and miR-145) and its calculating formula was as follows:  $\text{logit}(P) = 0.1559 \times \text{miR-15b} - 0.2063 \times \text{miR-17} + 0.0712 \times \text{miR-21} + 0.2252 \times \text{miR-26b} + 0.1311 \times \text{miR-145}$ . The other consists of the above 5 miRNAs plus one tumor marker (CEA), and its calculating formula was as follows:  $\text{logit}(P) = 0.5640 \times \text{CEA} + 0.1758 \times \text{miR-15b} - 0.2653 \times \text{miR-17} + 0.0466 \times \text{miR-21} + 0.1871 \times \text{miR-26b} + 0.2369 \times \text{miR-145}$ . As shown in Figure 4B, the ROC curves demonstrated that the combination of 5 miRNAs plus CEA has a better discriminatory performance, with an AUC of 0.85, when compared with the 5 miRNAs (AUC: 0.681) or CEA alone (AUC: 0.793).

### Diagnostic model confirmation in the validation phase

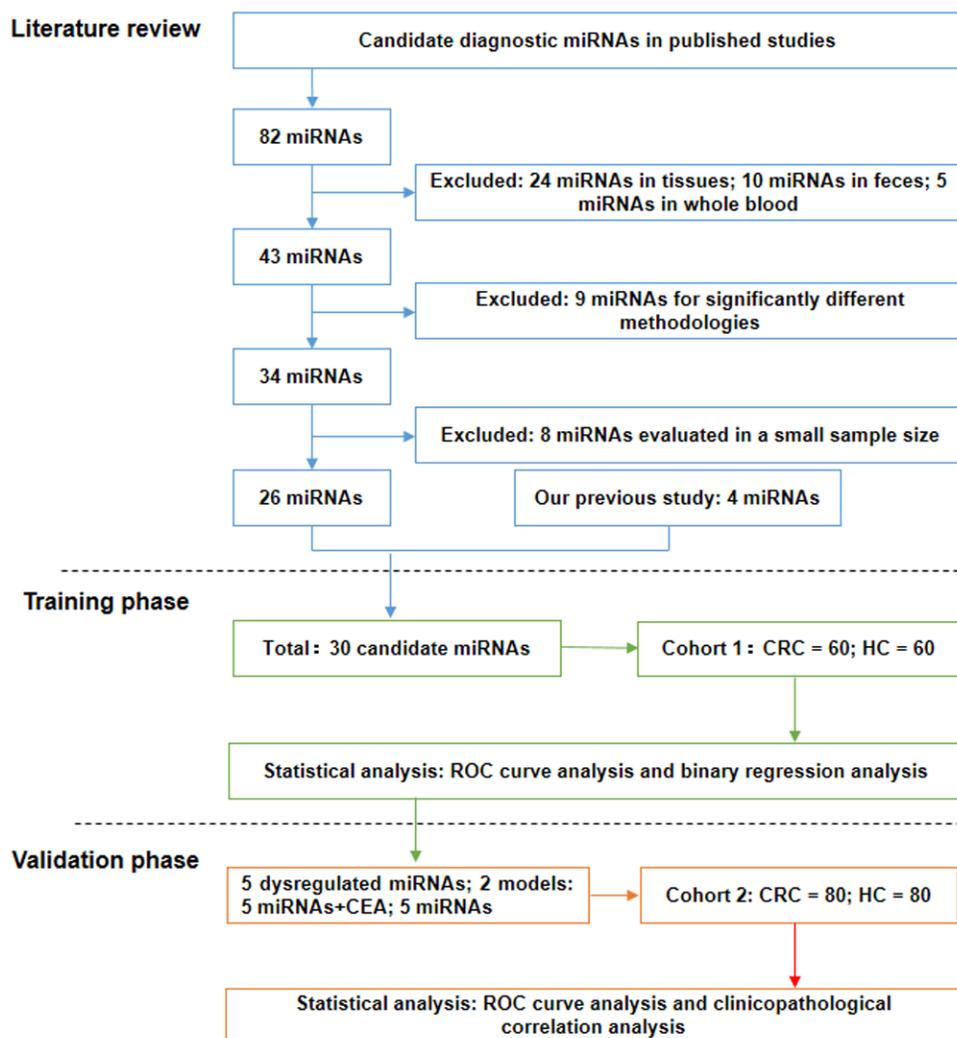
To confirm the discriminatory capability of the 5 dysregulated miRNAs and the diagnostic models detected in the training phase, an independent cohort of 80 CRC patients

and 80 HC was utilized. As shown in Figure 3, we found by qRT-PCR analysis that all the dysregulated miRNAs exhibited similar expression patterns in the training phase (miR-15b,  $p = 0.0071$ ; miR-17,  $p = 0.0005$ ; miR-21,  $p = 0.0011$ ; miR-26b,  $p < 0.0001$ ; miR-145,  $p = 0.0047$ ).

Roc curves were subsequently performed to validate the discriminatory performance of these miRNAs. MiR-26b had the best discriminatory performance (AUC: 0.708), while miR-15b had the worst performance (AUC: 0.624; Figure 4C and Table 2). Using the same formulas we further examined the efficiency of the diagnostic models constructed in the training cohort. As shown in Figure 4D, the model integrating 5 miRNAs and CEA had better discriminatory performance (AUC: 0.818) than the models comprising only the 5 miRNAs (AUC: 0.666) or just CEA (AUC: 0.790).

## Correlation between miRNAs and clinicopathological parameters in CRC

A summary of the correlations between the 5 dysregulated miRNAs identified above and the clinicopathological parameters of all the enrolled CRC patients ( $n = 140$ ) is presented in Supplementary Table 2. MiR-26b expression was significantly correlated with cancer type ( $p = 0.018$ ) and tumor size ( $p = 0.004$ ), while miR-145 expression was significantly correlated with tumor size only ( $p = 0.047$ ). No correlations were found between other circulating miRNA levels and clinicopathological parameters including age, gender, T status, N status, M status, TNM stage, tumor differentiation, tumor size, tumor location, cancer type, and CEA levels.



**Figure 1: Study design flow chart.** The study consisted of three sections: a systematic literature review to select candidate miRNAs, a training phase for screening candidate miRNAs and constructing diagnostic models, and a validation phase for confirming optimal miRNAs/diagnostic models.

**Table 1: Clinicopathological characteristics of all the participants in the training and validation cohort**

Characteristics	Training cohort		Validation cohort	
	CRC cases	Healthy controls	CRC cases	Healthy controls
	N=60	N=60	N=80	N=80
Age(Mean±SD,years)	61.20±11.00	60.98±5.26	63.75±12.34	62.25±8.24
Gender				
Male	36 (60%)	31 (51.7%)	50 (62.5%)	47 (58.75%)
Female	24 (40%)	29 (48.3%)	30 (37.5%)	33 (41.25%)
T status				
T1	7	-	4	-
T2	5	-	5	-
T3	7	-	7	-
T4	41	-	59	-
Unknown	0		5	
N status				
N0	29	-	35	-
N1	16	-	19	-
N2	15	-	21	-
Unknown	0		5	
M status				
M0	53	-	64	-
M1	7	-	11	-
Unknown	0		5	
UICC TNM stage				
I	12	-	8	-
II	17	-	27	-
III	24	-	29	-
IV	7	-	11	-
Unknown	0		5	
Tumor differentiation				
Well	7	-	8	-
Moderate	40	-	53	-
Poor	13	-	12	-
Unknown	0		7	
Tumor size (mm)				
≥5	33	-	40	-
<5	27	-	35	-
Unknown	0		5	
Tumor location				
Right sided colon	16	-	22	-

*(continued)*

Characteristics	Training cohort		Validation cohort	
	CRC cases	Healthy controls	CRC cases	Healthy controls
	N=60	N=60	N=80	N=80
Left sided colon	15	-	28	-
Rectum	29	-	30	-
CEA level				
Low (<5ng/ml)	28	60	42	80
High (≥5ng/ml)	29	0	34	0
Unknown	3	0	4	0
CA 19-9 level				
Low (<27U/ml)	39	57	54	68
High (≥27U/ml)	17	1	21	3
Unknown	4	2	5	9
CA125 level				
Low (<35U/ml)	51	58	61	77
High (≥35U/ml)	5	0	12	1
Unknown	4	2	7	2
CA-153 level				
Low (<25U/ml)	52	55	64	69
High (≥25U/ml)	0	1	4	0
Unknown	8	4	12	11
CA-724 level				
Low (<6.9U/ml)	38	58	42	79
High (≥6.9U/ml)	10	0	17	1
Unknown	12	2	21	0

## DISCUSSION

Tumor-derived circulating miRNAs have attracted wide attention in the field of oncological diagnosis since Mitchell et al. first proposed that they could be stable biomarkers for blood-based detection of solid cancers [41, 42]. Although numerous studies have extensively investigated miRNAs as potential diagnostic markers for CRC, most of them failed to gain acceptance in clinical practice mainly due to insufficient validation, unstandardized methodologies, and lack of cross-validation in terms of population ethnicity [43, 44]. Therefore, identifying studies that can help improve diagnostic efficiency is of great clinical significance. In the present study, we performed a systemic literature review to select candidate circulating miRNAs previously identified in different study cohorts. We additionally enrolled 4 miRNAs found to be associated with malignant characteristics of CRC in our previous studies, thus

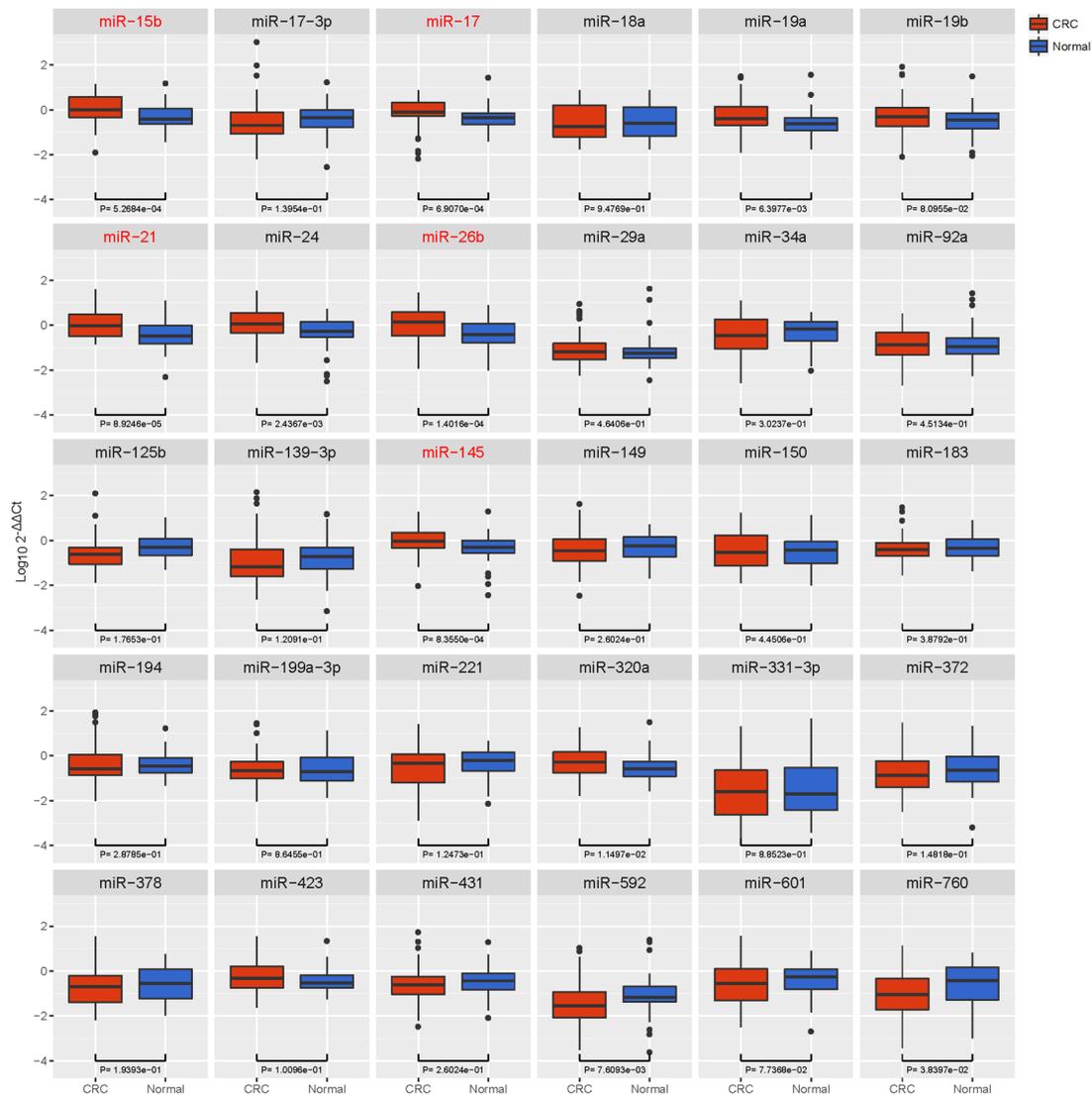
defining a total of 30 candidate miRNAs for further screening.

In the training phase, we screened out 5 circulating miRNAs (miR-15b, miR-17, miR-21, miR-26b, and miR-145) that proved to be closely linked to the initiation and development of disease in CRC patients. For instance, miR-15b was suggested to promote the invasion and metastasis of CRC cells, while a protective effect was described instead for miR-145 [45, 46]. On the other hand, both miR-21 and miR-26b were identified as crucial drivers of colitis-associated carcinogenesis [9, 47], whereas miR-17 was shown to contribute to CRC progression by inducing epithelial-mesenchymal transition and cancer stem cell phenotype [48]. ROC analysis further demonstrated that miR-21 had the best performance in discriminating CRC patients from HC among these 5 miRNAs. This observation is consistent with several recent studies supporting circulating miR-21 as an effective CRC biomarker [39, 49, 50], although there

are reports that question its practical utility in diagnosing CRC [51, 52]. As for miR-15b, miR-17, and miR-145, our results also agree with previous studies, and to our knowledge no related divergent reports were published [21, 38, 52]. Then, we employed a validation cohort to confirm the expression patterns of the 5 miRNAs selected in the training phase. Although these miRNAs remained significantly dysregulated in CRC patients, ROC analysis indicated that miR-26b has the best discriminating performance, which is somewhat inconsistent with our observation in the training phase. A recent work by Cristóbal et al. suggested that miR-26b overexpression might be correlated with lung metastasis in CRC [53].

Although to our knowledge there are currently no other available studies addressing circulating miR-26b levels in CRC patients, it should be mentioned that it had a stable and comparatively good discriminating performance both in the training and validation phases in our study, with an AUC of 0.702 and 0.708 respectively. Therefore, we conclude that miR-26b may be an effective plasma marker for CRC diagnosis.

Perhaps the most remarkable finding of our research was the construction and validation of a new diagnostic model integrating multiple miRNAs and CEA, a recognized CRC marker. Using logistic regression analysis, we found that our first diagnostic model, which



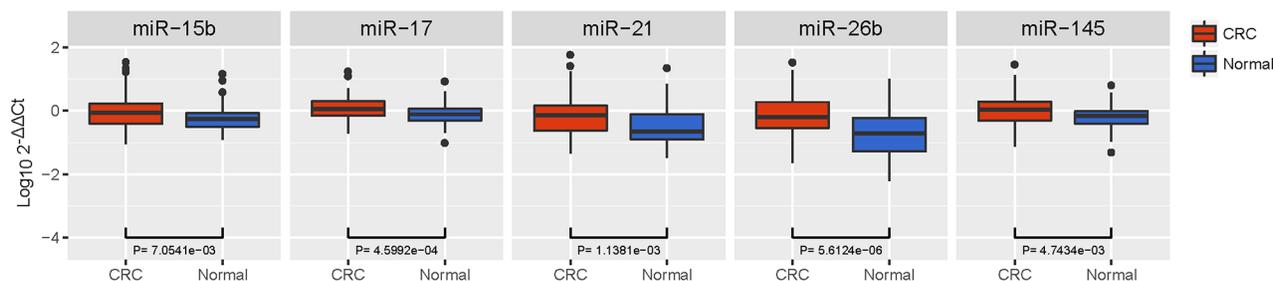
**Figure 2: Box plot of the relative expression levels of candidate miRNAs in the training cohort (n = 120).** The y axis indicates  $\log_{10} 2^{-\Delta\Delta C_t}$  relative expression of miRNA. Among all the circulating miRNAs selected (n = 30), the expression of 5 miRNAs was significantly higher in CRC than in HC (miR-15b, p = 0.0005; miR-17, p = 0.0007; miR-21, p < 0.0001; miR-26b, p = 0.0001; miR-145, p = 0.0008).

**Table 2: Diagnostic performances of the 5 dysregulated circulating miRNAs in the training and validation cohort**

Study cohort	miRNAs	Fold change	AUC(95%CI)	P value	Cut-off	Sensitivity(95%CI)	Specificity(95%CI)
Training	miR-15b	2.550	0.684(0.587-0.780)	0.0005	0.527	70.00%(56.79%-81.16%)	65.00%(51.60%-76.87%)
	miR-17	1.646	0.680(0.580-0.780)	0.0007	0.971	85.00%(73.43%-92.90%)	45.00%(32.12%-58.39%)
	miR-21	3.135	0.708(0.616-0.799)	<0.0001	0.686	71.67%(58.56%-82.55%)	58.33%(44.88%-70.93%)
	miR-26b	3.434	0.702(0.608-0.795)	0.0001	0.530	61.67%(48.21%-73.93%)	70.00%(56.79%-81.16%)
	miR-145	2.145	0.677(0.581-0.773)	0.0008	0.840	68.33%(55.04%-79.74%)	53.33%(40.00%-66.33%)
Validation	miR-15b	2.659	0.624(0.535-0.712)	0.0071	0.527	50.00%(38.61%-61.39%)	68.75%(57.41%-78.65%)
	miR-17	1.767	0.660(0.576-0.745)	0.0005	0.971	67.50%(56.11%-77.55%)	62.50%(50.96%-73.08%)
	miR-21	2.802	0.649(0.563-0.735)	0.0011	0.686	71.25%(60.05%-80.82%)	52.50%(41.02%-63.79%)
	miR-26b	2.796	0.708(0.627-0.789)	<0.0001	0.530	72.50%(61.38%-81.90%)	56.25%(44.70%-67.32%)
	miR-145	2.488	0.629(0.542-0.716)	0.0047	0.840	62.50%(50.96%-73.08%)	61.25%(49.70%-71.94%)

combined the 5 screened miRNAs (miR-15b, miR-17, miR-21, miR-26b, and miR-145) had an inferior diagnostic performance as compared with the single most optimal miRNA (AUC in the training phase: 0.681 vs 0.708; AUC in the validation phase: 0.666 vs 0.708). This result is in accordance with a recent meta-analysis that suggested that a single circulating miRNA, miR-21, performed significantly better than various circulating miRNA panels in CRC diagnosis [54]. However, another meta-analysis proposed that multiple circulating miRNAs could dramatically improve diagnostic accuracy compared with individual ones, with AUC values ranging from 0.79 to 0.89 [55]. These divergent observations can be

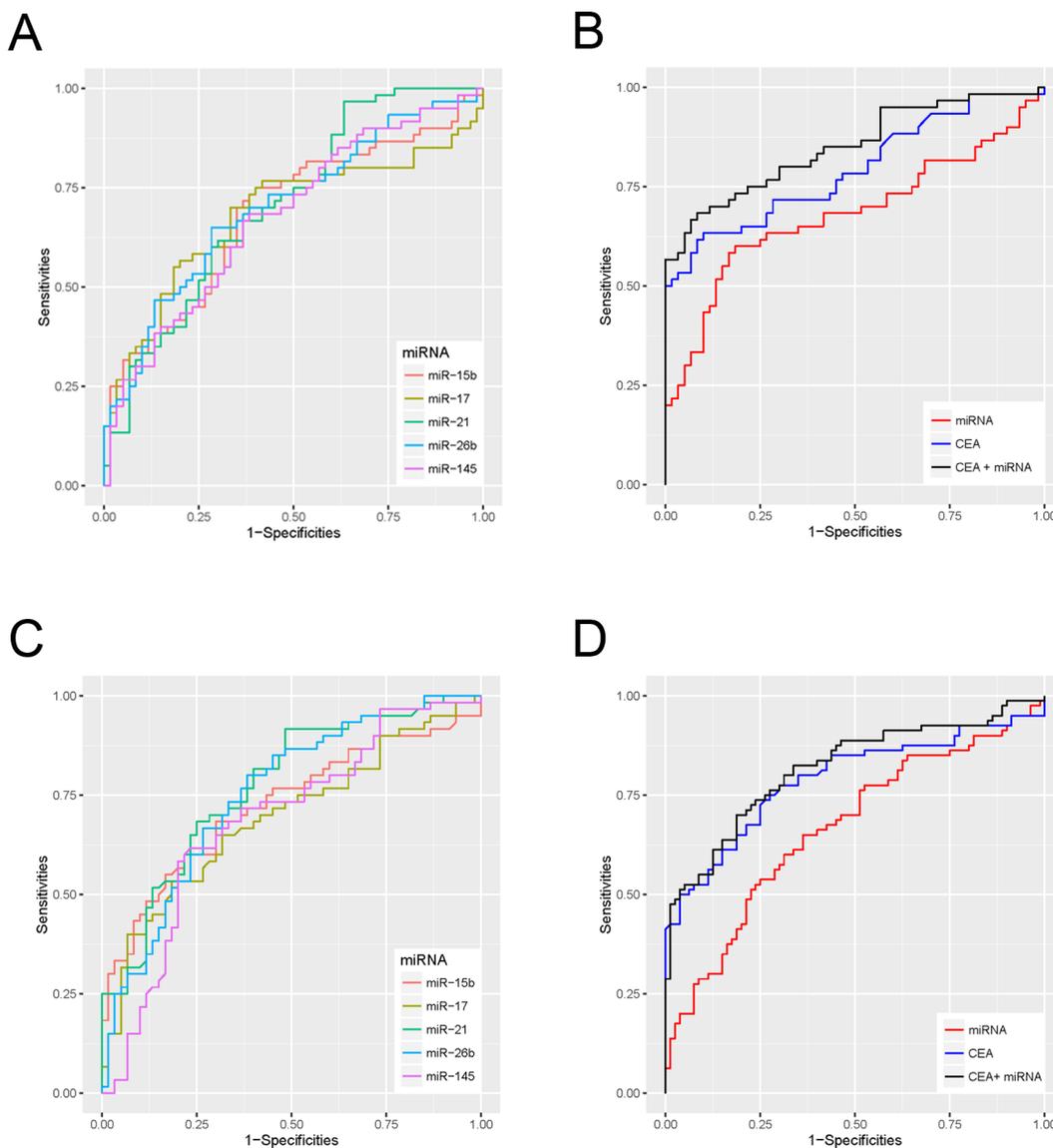
partly attributed to the heterogeneous expression of the several circulating miRNAs detected in CRC patients, and remain to be confirmed by further standardized clinical validations. We next compared the 5 miRNAs with CEA level and found that the miRNA panel had a significantly worse discriminating performance both in the training and validation phases. Although an opposite conclusion was reached by a previous study comparing multiple miRNAs with CEA level [20], our data thus suggests that the diagnostic efficiency of this miRNA combination set is inferior to current CEA detection. However, once integrated with CEA, our miRNA panel significantly improved the AUC for CEA, from 0.793 to 0.85 in the



**Figure 3: Box plot of the relative expression levels of dysregulated miRNAs in the validation cohort (n = 160).** The y axis indicates  $\log_{10} 2^{-\Delta\Delta C_t}$  relative expression of miRNA. All 5 dysregulated miRNAs exhibited similar expression patterns in the training phase (miR-15b, p = 0.0071; miR-17, p = 0.0005; miR-21, p = 0.0011; miR-26b, p < 0.0001; miR-145, p = 0.0047).

training phase, and from 0.790 to 0.818 in the validation phase. Although the validity of CEA determination in CRC diagnosis has been disputed, it remains the most commonly examined tumor marker for noninvasive diagnosis due to its high CRC specificity. The present results suggest that a novel model combining multiple circulating miRNAs and CEA level detection improves diagnostic efficiency in CRC, and could therefore be adopted successfully in the clinical practice. Furthermore, this approach appears to be more convenient and acceptable for patients in the diagnostic process, compared to conventional stool test and colonoscopy.

Despite its novel findings, our study has some potential deficiencies that are worth considering. Firstly, the majority of samples were obtained from stage III/IV patients, which prevented identifying eligible circulating miRNAs related to early stage CRC. This deficiency is largely attributed to the fact that most CRC patients are diagnosed at an advanced stage, while much fewer patients are diagnosed at stage I/II. To tackle this issue, more multicenter collaborations backed by appropriate clinical resources are strongly advocated. Secondly, in our study, combined miRNAs is inferior to single miR-21 or miR-26b in diagnostic performances, which is inconsistent



**Figure 4: Receiver Operating Characteristic (ROC) curves for dysregulated miRNAs and diagnostic models in the training and the validation cohort. (A)** ROC curves for the diagnostic performances of dysregulated miRNAs in the training cohort. **(B)** ROC curves for the diagnostic performances of the 5 miRNAs + CEA, the 5 miRNAs alone, and CEA alone in the training cohort. **(C)** ROC curves for the diagnostic performance of the 5 dysregulated miRNAs in the validation cohort. **(D)** ROC curves for the diagnostic performances of 5 miRNAs + CEA, 5 miRNAs alone, and CEA alone in the validation cohort.

with some studies. This issue is expected to be solved by further highly standardized validations in future. Thirdly, the diagnostic performance of our model is slightly inferior to that of other blood-based diagnostic markers such as methylated SPET9 (miRNAs + CEA, AUC: 0.85 (training)/0.818(validation) vs methylated SPET9, AUC: 0.88) [56]. In this regard, clinical validation is suggested in order to examine the efficacy and improve the construction of our model. On the other hand, additional efforts can be made to integrate miRNAs with other types of diagnostic biomarkers to create a more effective diagnostic or screening system. Finally, due to insufficient sampling we failed to identify dysregulated circulating miRNAs in patients with precancerous lesions such as advanced adenoma or ulcerative colitis. Therefore, their clinical utility in assessing CRC risk through population screening is worthy of further inquiry.

In conclusion, through systematic literature revision and clinical validation, our study identified 5 miRNAs differentially dysregulated in the plasma of CRC patients. By integrating these miRNAs and CEA level, we constructed and validated a new model that we believe will enhance the diagnostic accuracy of CRC and overcome some limitations of current blood-based diagnostic methods.

## MATERIALS AND METHODS

### Study design and miRNA selection

A flow chart of this study is illustrated in Figure 1. The study consisted of three general parts: a systematic literature review for selecting candidate miRNAs, a training phase for screening candidate miRNAs and constructing diagnostic models, and a validation phase for confirming optimal miRNAs/diagnostic models. For the literature review, we preliminarily selected candidate miRNAs from published studies based on the following inclusion criteria: diagnostic potential confirmed by at least 2 publications or CRC-related miRNAs identified in our previous studies. Then, we excluded unqualified candidates according to the following exclusion criteria: 1) miRNAs detected in tissue, feces, or whole blood; 2) obvious differences in methodology; 3) miRNAs detected in small samples ( $n \leq 50$ ). Next, the remaining candidates were screened in the training phase using qRT-PCR and integrated with traditional tumor markers (CEA, CA19-9, CA-724, CA153, and CA125) in diagnostic models. In the validation phase, the optimal diagnostic miRNAs and integrated models were examined in an independent cohort.

### Patient data and sample preparation

Between January 11, 2014 and April 8, 2016, a total of 60 newly diagnosed CRC patients and 60 healthy volunteers were enrolled from Shanghai Tenth People's Hospital Affiliated to Tongji University and allocated

into a training cohort. In addition, a validation cohort including 80 newly diagnosed CRC patients and 80 healthy volunteers was enrolled from The Sixth People's Hospital Affiliated to Shanghai Jiao Tong University between July 6, 2011, and August 23, 2012. All the enrolled cases were clinicopathologically confirmed as CRC. None of the patients had family history of cancer nor received previous chemoradiotherapy treatment. Tumor Node Metastasis stage was classified according to the 7<sup>th</sup> Union for International Cancer Control guidelines. The basic clinicopathological features of the two cohorts are presented in Table 1. This study was approved by the ethics committees of both hospitals and informed consent was obtained from all participants.

For sample preparation, 4 ml of peripheral blood was collected from each participant and transferred into EDTA tubes. Blood samples were centrifuged at 3000 rpm for 10 min and the supernatants were collected and stored at  $-80^{\circ}\text{C}$  before further processing.

### RNA extraction and quantitative real-time PCR

Total RNA was extracted from plasma samples using an RNA isolation kit (Qiagen, Hilden, Germany) following the manufacturer's instructions. Briefly, 200  $\mu\text{l}$  of plasma was thawed on ice, mixed with 1 ml QIAzol Lysis Reagent, and incubated at room temperature for 5 min. Inter-sample variation during RNA extraction was normalized using synthetic *C. elegans* cel-miR-39 ( $1.6 \times 10^8$  copies/ $\mu\text{l}$ ). Finally, the concentration of RNA samples was quantified on a NanoDrop ND-1000 spectrophotometer (Nano Drop Technologies, Wilmington, DE, USA).

The RNA thus obtained was reversely transcribed into cDNA using a MicroRNA Reverse Transcription Kit (BioTNT, Shanghai, China) according to the manufacturer's instructions. qRT-PCR was performed on a ViiA<sup>TM</sup> 7 Real-Time PCR System (Life Technologies, USA) under the following conditions:  $95^{\circ}\text{C}$  for 5 min, followed by 40 cycles at  $95^{\circ}\text{C}$  for 5 s, and  $60^{\circ}$  for 30 s. The reaction mixtures included 1  $\mu\text{l}$  cDNA, 5  $\mu\text{l}$  2 $\times$ qPCR Premix, 1  $\mu\text{l}$  microRNA upper primer and lower primer, and 2  $\mu\text{l}$  RNase-free water. The specificity and identity of the reaction products were verified by dissociation curve analysis. Data normalization was conducted using an exogenous (cel-miR-39) and an endogenous (miR-16-5p) control as described previously [57]. The relative level of each miRNA was calculated using the  $2^{-\Delta\Delta\text{Ct}}$  method and all the assays were carried out in triplicate. Information of the primers used for miRNA amplification is supplied in Supplementary Table 3.

### Statistical analysis

Data are presented as mean  $\pm$  SD. The baseline clinicopathological characteristics of each cohort were

compared by student t test or chi-square test. Relative miRNAs levels in the CRC and HC groups were compared by Mann-Whitney test and two-sided  $\alpha$  level was adjusted by simple Bonferroni correction. Correlations between miRNA levels and clinicopathological features were analyzed by Mann-Whitney tests. ROC curves and AUC values were used to describe the diagnostic performance of miRNAs and diagnostic models. The diagnostic models were constructed using binary logistic regression analysis. In brief, an accessing language procedure was used to integrate the miRNAs and serum tumor markers. A scoring formula was then established by assigning the coefficient to each included variable and the score of each patient was calculated accordingly. All the statistical analyses were performed using R version 3.2.3. A  $p < 0.05$  was considered statistically significant.

## ACKNOWLEDGMENTS

This study was funded by the National High Technology Research and Development Program (863 Program; grant number 2014AA020803), the National Natural Science Foundation of China (grant number 81302066), and Tongji University key construction program (No. 162385). We deeply appreciate the crucial cooperation of all the cancer registries regarding clinical data collection and validation.

## CONFLICTS OF INTEREST

None.

## REFERENCES

1. Siegel R, DeSantis C, Jemal A. Colorectal cancer statistics, 2014. *CA Cancer J Clin.* 2014; 64: 104-17. <https://doi.org/10.3322/caac.21220>.
2. Thomas DS, Fourkala EO, Apostolidou S, Gunu R, Ryan A, Jacobs I, Menon U, Alderton W, Gentry-Maharaj A, Timms JF. Evaluation of serum CEA, CYFRA21-1 and CA125 for the early detection of colorectal cancer using longitudinal preclinical samples. *Br J Cancer.* 2015; 113: 268-74. <https://doi.org/10.1038/bjc.2015.202>.
3. Aghagolzadeh P, Radpour R. New trends in molecular and cellular biomarker discovery for colorectal cancer. *World J Gastroenterol.* 2016; 22: 5678-93. <https://doi.org/10.3748/wjg.v22.i25.5678>.
4. Svoronos AA, Engelman DM, Slack FJ. OncomiR or tumor suppressor? The duplicity of MicroRNAs in cancer. *Cancer Res.* 2016; 76: 3666-70. <https://doi.org/10.1158/0008-5472.CAN-16-0359>.
5. Ma Y, Zhang P, Wang F, Zhang H, Yang Y, Shi C, Xia Y, Peng J, Liu W, Yang Z, Qin H. Elevated oncofetal miR-17-5p expression regulates colorectal cancer progression by repressing its target gene P130. *Nat Commun.* 2012; 3: 1291. <https://doi.org/10.1038/ncomms2276>.
6. Wang F, Ma YL, Zhang P, Shen TY, Shi CZ, Yang YZ, Moyer MP, Zhang HZ, Chen HQ, Liang Y, Qin HL. SP1 mediates the link between methylation of the tumour suppressor miR-149 and outcome in colorectal cancer. *J Pathol.* 2013; 229: 12-24. <https://doi.org/10.1002/path.4078>.
7. Ma Y, Zhang P, Wang F, Zhang H, Yang J, Peng J, Liu W, Qin H. miR-150 as a potential biomarker associated with prognosis and therapeutic outcome in colorectal cancer. *Gut.* 2012; 61: 1447-53. <https://doi.org/10.1136/gutjnl-2011-301122>.
8. Feng J, Yang Y, Zhang P, Wang F, Ma Y, Qin H, Wang Y. miR-150 functions as a tumour suppressor in human colorectal cancer by targeting c-Myb. *J Cell Mol Med.* 2014; 18: 2125-34. <https://doi.org/10.1111/jcmm.12398>.
9. Shi C, Yang Y, Xia Y, Okugawa Y, Yang J, Liang Y, Chen H, Zhang P, Wang F, Han H, Wu W, Gao R, Gasche C, et al. Novel evidence for an oncogenic role of microRNA-21 in colitis-associated colorectal cancer. *Gut.* 2016; 65: 1470-80. <https://doi.org/10.1136/gutjnl-2014-308455>.
10. Weber JA, Baxter DH, Zhang S, Huang DY, Huang KH, Lee MJ, Galas DJ, Wang K. The microRNA spectrum in 12 body fluids. *Clin Chem.* 2010; 56: 1733-41. <https://doi.org/10.1373/clinchem.2010.147405>.
11. Xiao YF, Yong X, Fan YH, Lu MH, Yang SM, Hu CJ. microRNA detection in feces, sputum, pleural effusion and urine: novel tools for cancer screening (Review). *Oncol Rep.* 2013; 30: 535-44. <https://doi.org/10.3892/or.2013.2525>.
12. Ng EK, Chong WW, Jin H, Lam EK, Shin VY, Yu J, Poon TC, Ng SS, Sung JJ. Differential expression of microRNAs in plasma of patients with colorectal cancer: a potential marker for colorectal cancer screening. *Gut.* 2009; 58: 1375-81. <https://doi.org/10.1136/gut.2008.167817>.
13. Kanaan Z, Rai SN, Eichenberger MR, Roberts H, Keskey B, Pan J, Galandiuk S. Plasma miR-21: a potential diagnostic marker of colorectal cancer. *Ann Surg.* 2012; 256: 544-51. <https://doi.org/10.1097/SLA.0b013e318265bd6f>.
14. Liu GH, Zhou ZG, Chen R, Wang MJ, Zhou B, Li Y, Sun XF. Serum miR-21 and miR-92a as biomarkers in the diagnosis and prognosis of colorectal cancer. *Tumour Biol.* 2013; 34: 2175-81. <https://doi.org/10.1007/s13277-013-0753-8>.
15. Pu XX, Huang GL, Guo HQ, Guo CC, Li H, Ye S, Ling S, Jiang L, Tian Y, Lin TY. Circulating miR-221 directly amplified from plasma is a potential diagnostic and prognostic marker of colorectal cancer and is correlated with p53 expression. *J Gastroenterol Hepatol.* 2010; 25: 1674-80. <https://doi.org/10.1111/j.1440-1746.2010.06417.x>.
16. Cheng G. Circulating miRNAs: roles in cancer diagnosis, prognosis and therapy. *Adv Drug Deliv Rev.* 2015; 81: 75-93. <https://doi.org/10.1016/j.addr.2014.09.001>.
17. Pritchard CC, Kroh E, Wood B, Arroyo JD, Dougherty KJ, Miyaji MM, Tait JF, Tewari M. Blood cell origin of

- circulating microRNAs: a cautionary note for cancer biomarker studies. *Cancer Prev Res (Phila)*. 2012; 5: 492-7. <https://doi.org/10.1158/1940-6207.CAPR-11-0370>.
18. Toiyama Y, Takahashi M, Hur K, Nagasaka T, Tanaka K, Inoue Y, Kusunoki M, Boland CR, Goel A. Serum miR-21 as a diagnostic and prognostic biomarker in colorectal cancer. *J Natl Cancer Inst*. 2013; 105: 849-59. <https://doi.org/10.1093/jnci/djt101>.
  19. Fang Z, Tang J, Bai Y, Lin H, You H, Jin H, Lin L, You P, Li J, Dai Z, Liang X, Su Y, Hu Q, et al. Plasma levels of microRNA-24, microRNA-320a, and microRNA-423-5p are potential biomarkers for colorectal carcinoma. *J Exp Clin Cancer Res*. 2015; 34: 86. <https://doi.org/10.1186/s13046-015-0198-6>.
  20. Wang J, Huang SK, Zhao M, Yang M, Zhong JL, Gu YY, Peng H, Che YQ, Huang CZ. Identification of a circulating microRNA signature for colorectal cancer detection. *PLoS One*. 2014; 9: e87451. <https://doi.org/10.1371/journal.pone.0087451>.
  21. Giraldez MD, Lozano JJ, Ramirez G, Hijona E, Bujanda L, Castells A, Gironella M. Circulating microRNAs as biomarkers of colorectal cancer: results from a genome-wide profiling and validation study. *Clin Gastroenterol Hepatol*. 2013; 11: 681-8. <https://doi.org/10.1016/j.cgh.2012.12.009>.
  22. Kanaan Z, Roberts H, Eichenberger MR, Billeter A, Ocheretner G, Pan J, Rai SN, Jordan J, Williford A, Galandiuk S. A plasma microRNA panel for detection of colorectal adenomas: a step toward more precise screening for colorectal cancer. *Ann Surg*. 2013; 258: 400-8. <https://doi.org/10.1097/SLA.0b013e3182a15bcc>.
  23. Faltejskova P, Bocanek O, Sachlova M, Svoboda M, Kiss I, Vyzula R, Slaby O. Circulating miR-17-3p, miR-29a, miR-92a and miR-135b in serum: evidence against their usage as biomarkers in colorectal cancer. *Cancer Biomark*. 2012; 12: 199-204. <https://doi.org/10.3233/CBM-130308>.
  24. 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: 320-6. <https://doi.org/10.3892/or.2013.2475>.
  25. Wang B, Zhang Q. The expression and clinical significance of circulating microRNA-21 in serum of five solid tumors. *J Cancer Res Clin Oncol*. 2012; 138: 1659-66. <https://doi.org/10.1007/s00432-012-1244-9>.
  26. Nonaka R, Nishimura J, Kagawa Y, Osawa H, Hasegawa J, Murata K, Okamura S, Ota H, Uemura M, Hata T, Takemasa I, Mizushima T, Okuzaki D, et al. Circulating miR-199a-3p as a novel serum biomarker for colorectal cancer. *Oncol Rep*. 2014; 32: 2354-8. <https://doi.org/10.3892/or.2014.3515>.
  27. Basati G, Emami Razavi A, Abdi S, Mirzaei A. Elevated level of microRNA-21 in the serum of patients with colorectal cancer. *Med Oncol*. 2014; 31: 205. <https://doi.org/10.1007/s12032-014-0205-3>.
  28. Zanutto S, Pizzamiglio S, Ghilotti M, Bertan C, Ravagnani F, Perrone F, Leo E, Pilotti S, Verderio P, Gariboldi M, Pierotti MA. Circulating miR-378 in plasma: a reliable, haemolysis-independent biomarker for colorectal cancer. *Br J Cancer*. 2014; 110: 1001-7. <https://doi.org/10.1038/bjc.2013.819>.
  29. Du M, Liu S, Gu D, Wang Q, Zhu L, Kang M, Shi D, Chu H, Tong N, Chen J, Adams TS, Zhang Z, Wang M. Clinical potential role of circulating microRNAs in early diagnosis of colorectal cancer patients. *Carcinogenesis*. 2014; 35: 2723-30. <https://doi.org/10.1093/carcin/bgu189>.
  30. Huang Z, Huang D, Ni S, Peng Z, Sheng W, Du X. Plasma microRNAs are promising novel biomarkers for early detection of colorectal cancer. *Int J Cancer*. 2010; 127: 118-26. <https://doi.org/10.1002/ijc.25007>.
  31. Hofslie E, Sjursen W, Prestvik WS, Johansen J, Rye M, Trano G, Wasmuth HH, Hatlevoll I, Thommesen L. Identification of serum microRNA profiles in colon cancer. *Br J Cancer*. 2013; 108: 1712-9. <https://doi.org/10.1038/bjc.2013.121>.
  32. Aherne ST, Madden SF, Hughes DJ, Pardini B, Naccarati A, Levy M, Vodicka P, Neary P, Dowling P, Clynes M. Circulating miRNAs miR-34a and miR-150 associated with colorectal cancer progression. *BMC Cancer*. 2015; 15: 329. <https://doi.org/10.1186/s12885-015-1327-5>.
  33. Yuan D, Li K, Zhu K, Yan R, Dang C. Plasma miR-183 predicts recurrence and prognosis in patients with colorectal cancer. *Cancer Biol Ther*. 2015; 16: 268-75. <https://doi.org/10.1080/15384047.2014.1002327>.
  34. Basati G, Razavi AE, Pakzad I, Malayeri FA. Circulating levels of the miRNAs, miR-194, and miR-29b, as clinically useful biomarkers for colorectal cancer. *Tumour Biol*. 2016; 37: 1781-8. <https://doi.org/10.1007/s13277-015-3967-0>.
  35. Yu J, Jin L, Jiang L, Gao L, Zhou J, Hu Y, Li W, Zhi Q, Zhu X. Serum miR-372 is a diagnostic and prognostic biomarker in patients with early colorectal cancer. *Anticancer Agents Med Chem*. 2016; 16: 424-31.
  36. Liu M, Zhi Q, Wang W, Zhang Q, Fang T, Ma Q. Up-regulation of miR-592 correlates with tumor progression and poor prognosis in patients with colorectal cancer. *Biomed Pharmacother*. 2015; 69: 214-20. <https://doi.org/10.1016/j.biopha.2014.12.001>.
  37. Wang Q, Huang Z, Ni S, Xiao X, Xu Q, Wang L, Huang D, Tan C, Sheng W, Du X. Plasma miR-601 and miR-760 are novel biomarkers for the early detection of colorectal cancer. *PLoS One*. 2012; 7: e44398. <https://doi.org/10.1371/journal.pone.0044398>.
  38. Luo X, Stock C, Burwinkel B, Brenner H. Identification and evaluation of plasma microRNAs for early detection of colorectal cancer. *PLoS One*. 2013; 8: e62880. <https://doi.org/10.1371/journal.pone.0062880>.
  39. Yamada A, Horimatsu T, Okugawa Y, Nishida N, Honjo H, Ida H, Kou T, Kusaka T, Sasaki Y, Yagi M, Higurashi T, Yukawa N, Amanuma Y, et al. Serum miR-21, miR-29a, and miR-125b are promising biomarkers for the early detection

- of colorectal neoplasia. *Clin Cancer Res.* 2015; 21: 4234-42. <https://doi.org/10.1158/1078-0432.CCR-14-2793>.
40. Ma YL, Zhang P, Wang F, Moyer MP, Yang JJ, Liu ZH, Peng JY, Chen HQ, Zhou YK, Liu WJ, Qin HL. Human embryonic stem cells and metastatic colorectal cancer cells shared the common endogenous human microRNA-26b. *J Cell Mol Med.* 2011; 15: 1941-54. <https://doi.org/10.1111/j.1582-4934.2010.01170.x>.
  41. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, Peterson A, Noteboom J, O'Briant KC, Allen A, Lin DW, Urban N, Drescher CW, et al. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci U S A.* 2008; 105: 10513-8. <https://doi.org/10.1073/pnas.0804549105>.
  42. Kawaguchi T, Komatsu S, Ichikawa D, Tsujiura M, Takeshita H, Hirajima S, Miyamae M, Okajima W, Ohashi T, Imamura T, Kiuchi J, Konishi H, Shiozaki A, et al. Circulating microRNAs: a next-generation clinical biomarker for digestive system cancers. *Int J Mol Sci.* 2016; 17: 1459. <https://doi.org/10.3390/ijms17091459>.
  43. Ren A, Dong Y, Tsoi H, Yu J. Detection of miRNA as non-invasive biomarkers of colorectal cancer. *Int J Mol Sci.* 2015; 16: 2810-23. <https://doi.org/10.3390/ijms16022810>.
  44. Yi R, Li Y, Wang FL, Miao G, Qi RM, Zhao YY. MicroRNAs as diagnostic and prognostic biomarkers in colorectal cancer. *World J Gastrointest Oncol.* 2016; 8: 330-40. <https://doi.org/10.4251/wjgo.v8.i4.330>.
  45. Li J, Chen Y, Guo X, Zhou L, Jia Z, Tang Y, Lin L, Liu W, Ren C. Inhibition of miR-15b decreases cell migration and metastasis in colorectal cancer. *Tumour Biol.* 2016; 37: 8765-73. <https://doi.org/10.1007/s13277-015-4396-9>.
  46. Wang W, Ji G, Xiao X, Chen X, Qin WW, Yang F, Li YF, Fan LN, Xi WJ, Huo Y, Wen WH, Yang AG, Wang T. Epigenetically regulated miR-145 suppresses colon cancer invasion and metastasis by targeting LASP1. *Oncotarget.* 2016; 7: 68674-87. <https://doi.org/10.18632/oncotarget.11919>.
  47. Benderska N, Dittrich AL, Knaup S, Rau TT, Neufert C, Wach S, Fahlbusch FB, Rauh M, Wirtz RM, Agaimy A, Srinivasan S, Mahadevan V, Rummele P, et al. miRNA-26b overexpression in ulcerative colitis-associated carcinogenesis. *Inflamm Bowel Dis.* 2015; 21: 2039-51. <https://doi.org/10.1097/MIB.0000000000000453>.
  48. Xi XP, Zhuang J, Teng MJ, Xia LJ, Yang MY, Liu QG, Chen JB. MicroRNA-17 induces epithelial-mesenchymal transition consistent with the cancer stem cell phenotype by regulating CYP7B1 expression in colon cancer. *Int J Mol Med.* 2016; 38: 499-506. <https://doi.org/10.3892/ijmm.2016.2624>.
  49. Chen H, Liu H, Zou H, Chen R, Dou Y, Sheng S, Dai S, Ai J, Melson J, Kittles RA, Pirooznia M, Liptay MJ, Borgia JA, et al. Evaluation of plasma miR-21 and miR-152 as diagnostic biomarkers for common types of human cancers. *J Cancer.* 2016; 7: 490-9. <https://doi.org/10.7150/jca.12351>.
  50. Carter JV, Roberts HL, Pan J, Rice JD, Burton JF, Galbraith NJ, Eichenberger MR, Jordan J, Deveaux P, Farmer R, Williford A, Kanaan Z, Rai SN, et al. A highly predictive model for diagnosis of colorectal neoplasms using plasma microRNA: improving specificity and sensitivity. *Ann Surg.* 2016; 264: 575-84. <https://doi.org/10.1097/SLA.0000000000001873>.
  51. Montagnana M, Benati M, Danese E, Minicozzi AM, Paviati E, Gusella M, Pasini F, Bovo C, Guidi GC, Lippi G. Plasma expression levels of circulating miR-21 are not useful for diagnosing and monitoring colorectal cancer. *Clin Lab.* 2016; 62: 967-70.
  52. Zekri AR, Youssef AS, Lotfy MM, Gabr R, Ahmed OS, Nassar A, Hussein N, Omran D, Medhat E, Eid S, Hussein MM, Ismail MY, Alenzi FQ, et al. Circulating serum miRNAs as diagnostic markers for colorectal cancer. *PLoS One.* 2016; 11: e0154130. <https://doi.org/10.1371/journal.pone.0154130>.
  53. Cristobal I, Manso R, Gonzalez-Alonso P, Madoz-Gurpide J, Rojo F, Garcia-Foncillas J. Clinical value of miR-26b discriminating ulcerative colitis-associated colorectal cancer in the subgroup of patients with metastatic disease. *Inflamm Bowel Dis.* 2015; 21: E24-5. <https://doi.org/10.1097/MIB.0000000000000572>.
  54. Rokkas T, Kothonas F, Rokka A, Koukoulis G, Symvoulakis E. The role of circulating microRNAs as novel biomarkers in diagnosing colorectal cancer: a meta-analysis. *Eur J Gastroenterol Hepatol.* 2015; 27: 819-25. <https://doi.org/10.1097/MEG.0000000000000363>.
  55. Wang R, Wen H, Xu Y, Chen Q, Luo Y, Lin Y, Luo Y, Xu A. Circulating microRNAs as a novel class of diagnostic biomarkers in gastrointestinal tumors detection: a meta-analysis based on 42 articles. *PLoS One.* 2014; 9: e113401. <https://doi.org/10.1371/journal.pone.0113401>.
  56. Nian J, Sun X, Ming S, Yan C, Ma Y, Feng Y, Yang L, Yu M, Zhang G, Wang X. Diagnostic accuracy of methylated SEPT9 for blood-based colorectal cancer detection: a systematic review and meta-analysis. *Clin Transl Gastroenterol.* 2017; 8: e216. <https://doi.org/10.1038/ctg.2016.66>.
  57. Zhang J, Song Y, Zhang C, Zhi X, Fu H, Ma Y, Chen Y, Pan F, Wang K, Ni J, Jin W, He X, Su H, et al. Circulating miR-16-5p and miR-19b-3p as two novel potential biomarkers to indicate progression of gastric cancer. *Theranostics.* 2015; 5: 733-45. <https://doi.org/10.7150/thno.10305>.