



# Integrated Analysis Identifies a Nine-microRNA Signature Biomarker for Diagnosis and Prognosis in Colorectal Cancer

Ziyang Di<sup>1,2</sup>, Maojun Di<sup>2</sup>, Weihua Fu<sup>1</sup>, Qiang Tang<sup>2</sup>, Yanwei Liu<sup>2</sup>, Peijie Lei<sup>3</sup>, Xincheng Gu<sup>4</sup>, Tong Liu<sup>1\*</sup> and Min Sun<sup>2,5,6\*</sup>

<sup>1</sup> Department of General Surgery, Tianjin Medical University General Hospital, Tianjin Medical University, Tianjin, China, <sup>2</sup> Department of General Surgery, Taihe Hospital, Hubei University of Medicine, Shiyan, China, <sup>3</sup> The First Clinical School, Hubei University of Medicine, Shiyan, China, <sup>4</sup> College of Basic Medical Sciences, Hubei University of Medicine, Shiyan, China, <sup>5</sup> Hubei Key Laboratory of Embryonic Stem Cell Research, Taihe Hospital, Hubei University of Medicine, Shiyan, China, <sup>6</sup> Department of Anesthesiology, Institute of Anesthesiology, Taihe Hospital, Hubei University of Medicine, Shiyan, China

## OPEN ACCESS

### Edited by:

Zhaohui Huang,  
Affiliated Hospital of Jiangnan  
University, China

### Reviewed by:

Meng Ding,  
Nanjing Drum Tower Hospital, China  
Olga Brovkina,  
Federal Medical-Biological Agency,  
Russia

### \*Correspondence:

Tong Liu  
liutonga@126.com  
Min Sun  
sunmin-0715@163.com

### Specialty section:

This article was submitted to  
Cancer Genetics,  
a section of the journal  
Frontiers in Genetics

Received: 08 October 2019

Accepted: 18 February 2020

Published: 20 March 2020

### Citation:

Di Z, Di M, Fu W, Tang Q, Liu Y,  
Lei P, Gu X, Liu T and Sun M (2020)  
Integrated Analysis Identifies  
a Nine-microRNA Signature  
Biomarker for Diagnosis  
and Prognosis in Colorectal Cancer.  
Front. Genet. 11:192.  
doi: 10.3389/fgene.2020.00192

**Background:** Colorectal cancer (CRC) is the third most lethal and malignant type of cancer in the world. Abnormal expression of human microRNA-200a (hsa-miRNA-200a or miR-200a) has previously been characterized as a clinically noticeable biomarker in several cancers, but its role in CRC is still unclear.

**Methods:** Three CRC miRNA expression datasets were integratively analyzed by Least Absolute Shrinkage and Selector Operation (LASSO) and Support Vector Machine-Recursive Feature Elimination (SVM-RFE) algorithms. Nine candidate miRNAs were identified and validated for diagnostic and prognostic capability with the prediction model. The potential roles of the tumor suppressor miR-200a-3p in invasion, migration, and epithelial-mesenchymal transition of CRC cells were elaborated by *in vitro* studies.

**Results:** Nine miRNAs (miR-492, miR-200a, miR-338, miR-29c, miR-101, miR-148a, miR-92a, miR-424, and miR-210) were identified as potentially useful diagnostic biomarkers in the clinic. The overall accuracy rate of the nine miRNAs in the diagnostic model was 0.94, 0.89, and 0.978 in the testing, validation, and independent validation dataset, respectively. CRC patients in the GSE29622 cohort were separated by the prognostic model into the low-risk score group and the high-risk score group. The area under the receiver operating characteristic curve (AUC) was 0.872 and 0.783 for predicting the 1- to 10-year survival of CRC patients. The performance of the prognostic model was validated by an independent TCGA-Colon Adenocarcinoma (COAD) dataset with AUC values between 0.911 and 0.796 in predicting 1- to 10-year survival. Nomograms comprising risk scores, tumor stage, and TNM staging were generated for predicting 1-, 3-, and 5-year overall survival (OS) in the GSE29622 and TCGA-COAD datasets. Colony formation, invasion, and migration in DLD1 and

SW480 cells were suppressed by overexpression of miR-200a-3p. Inhibition of miR-200a-3p function contributed to abnormal colony formation, migration, invasion, and epithelial–mesenchymal transition (EMT). miR-200a-3p binding sites were located within the 3′-untranslated region (3′-UTR) of the Forkhead box protein A1 (FOXA1) mRNA.

**Conclusion:** We developed and validated a diagnostic and prognostic prediction model for CRC. miR-200a-3p was determined to be a potential diagnostic and prognostic biomarker for CRC.

**Keywords:** hsa-miRNA-200a, FOXA1, colorectal cancer, prediction model, prognosis, biomarker

## INTRODUCTION

Colorectal cancer is the third most commonly diagnosed cancer in both men and women worldwide (Siegel et al., 2017). The 5-year survival percentage for localized stage CRC reaches 90%. However, as CRC cells are transferred to the regional lymph nodes or different organs of the body, the 5-year survival rate declines from 71 to 13% in the United States (Siegel et al., 2016). The prognosis of CRC patients is dependent on the TNM stage of CRC and curative surgical intervention, which is applicable only for patients with CRC limited to primary tumor and regional lymph nodes. Moreover, the clinical performance, therapeutic efficacy, and prognosis of the patients are affected by many factors such as the epigenetic status and microenvironments that cause CRC heterogeneity (Linnekamp et al., 2015). It is widely accepted that diagnosis and resection at a precancerous or early stage based on effective biomarkers are the keys to the most patients being cured (Ribic et al., 2003; Kennedy et al., 2011). Furthermore, it is imperative to identify biomarkers that are essential for accurately monitoring patients' progression and their response to treatment (Danielsen et al., 2018; Ozawa et al., 2018).

MicroRNAs (miRNAs) have been developed as novel biomarkers for early non-invasive or minimally invasive diagnosis of CRC (Gangadhar and Schilsky, 2010; Chen et al., 2015). Expression profiles of miRNAs in human colon cancer were utilized to determine the potential clinical relevance in several studies (Smits et al., 2011; Ma et al., 2012; Manceau et al., 2014; Wu et al., 2014; Kandimalla et al., 2018). A single model incorporating multiple biomarkers demonstrates improved performance and stability of prognostic value (Agesen et al., 2012; Marisa et al., 2013). However, further validation is still needed for the implementation of these miRNAs in standard CRC clinical practice, and the underlying mechanism remains to be elucidated (Huang et al., 2017; Lin et al., 2019).

**Abbreviations:** AUC, area under the receiver operating characteristic curve; COAD, colon Adenocarcinoma; CRC, colorectal cancer; DRS, diagnosis risk score; EMT, epithelial mesenchymal transition; FOXA1, forkhead box protein A1; FPKM, fragments per kilobase of transcript per Million mapped reads; GEO, gene expression omnibus; hsa-miRNA-200a or miR-200a, human microRNA-200a; LASSO, least absolute shrinkage and selector operation; NC, negative control; OS, overall survival; ANT, paired adjacent normal tissue; RIPA, radioimmunoprecipitation assay; ROC, receiver operating characteristic; SD, standard deviation; SVM, support vector machine; SVM-RFE, support vector machine-recursive feature elimination; TCGA, the cancer genome atlas; YAP, YES-associated protein.

The miR-200 family (miR-200a, -200b, -200c, -141, and -429) has been demonstrated as an inhibitor of epithelial-to-mesenchymal transition (EMT) and metastasis of cancer in colon (Hur et al., 2013). Chen et al. (2019) reported that miR-200a is a tumor suppressor in glioma by directly targeting forkhead box A1 (FOXA1). FOXA1 is a transcription factor that is essential for epithelial lineage differentiation, playing an oncogenic or tumor-suppressive role in human malignancies (Song et al., 2019). FOXA1 upregulates the expression of YAP, promoting CRC tumor growth (Ma et al., 2016). Our previous study confirmed that YAP mediates EMT (Sun et al., 2017a). FOXA1 might be the potential target of miR-200a, regulating YAP-mediated EMT; however, the underlying mechanism in CRC remains unclear.

In this four-phase study, we developed a data processing system to solve the curse of dimensionality in high-dimensional gene expression data using LASSO and SVM. Different independent datasets were first integrated by using Fisher's method to expand the sample size, and the integrated dataset was then screened for candidate miRNAs of CRC using a prediction model combining the LASSO and SVM models. The full-length 3′-UTR of human FOXA1 mRNA was proven for the first time to be a direct target of miR-200a-3p in CRC. A multi-miRNA-based classifier with a logistic regression model was developed for CRC screening or early diagnosis and was validated with the Cox regression model for potential predictors of prognosis. In addition, the results of the present study demonstrate a potential data processing model for identifying novel biomarkers and candidate miRNA patterns in the detection and prognosis prediction of CRC.

## MATERIALS AND METHODS

### Data Collection, Preprocessing, and Normalization

Public microarray datasets were extracted from the GEO and TCGA database. The checklist and pipeline for proper organization of the integrated analysis were determined following the reporting guidelines of microarray meta-analysis recommended by Ramasamy et al. (2008). Only original experimental studies to screen miRNAs that were differentially expressed (DE) between CRC and ANT in at least 40 human samples were included. Selection criteria, probe annotation,

and data normalization were the same as described in previous reports (Sun et al., 2017a,b; Lin et al., 2019).

## Integrated Analysis of miRNA Expression Datasets

Differentially expressed miRNAs between CRC and ANT were determined by MetaOmics software<sup>1</sup> in the MetaDE package (Wang et al., 2012). The filter thresholds of the mean and SD were set to 30% in integrated analysis. Fisher's method was performed for statistically significant analysis to counterpoise the different stringencies of the methods. A permutation method with a modified *t*-test was used for the removal of the batch effect and estimation of the *P*-values (Tseng et al., 2012). *P*-value < 0.001 was used as the cutoff for statistically significant DE miRNAs.

## Selection of miRNAs for the Training Cohort

The candidate miRNAs were determined by a LASSO algorithm with penalty parameter tuning with 10-fold cross-validation and an SVM-Recursive Feature Elimination (RFE) algorithm (Qiu et al., 2017). In total, nine common candidate miRNAs were further selected for the training cohort.

## Development and Validation of the Logistic Prediction Diagnosis Model

The diagnosis model was developed for evaluation of the differential capacity of CRC and ANT in GSE49246 using logistic regression of the nine candidate miRNAs. Furthermore, candidate miRNAs were used to validate the results of the ROC curve analysis in an independent validation set (GSE115513) by forecasting the grades of CRC and ANT, as described previously (Sun et al., 2017b).

## Prognostic Index of the Nine-miRNA Signature and Validation on Prognosis of CRC Survival

A nine-miRNA prognostic signature model was developed by multivariate Cox hazard model analysis of GSE29622 data to estimate prognostic risk score (PRS) using the formula:  $PRS = \sum(C \times EXP_{miRNA})$ , where EXP was the FPKM value of the miRNA, and C was the regression coefficient for the corresponding miRNA in multivariate Cox hazard model analysis. The median PRS of the testing dataset GSE29622 was used to distinguish the high-risk cluster from the low-risk cluster. The prognostic performance of the miRNA signature model was evaluated by comparing the AUC of ROC curves. The association of miRNA signature with patient survival was analyzed, and the miRNA signature model was finally validated in an independent dataset, TCGA-COAD.

## Target Prediction of miRNA

The speculated targets of integrated-signature miRNAs, particularly for the tumor suppressor miRNA hsa-miR-200a,

<sup>1</sup><http://www.pitt.edu/~tsengweb/MetaOmicsHome.htm>

were predicted by four different target-prediction algorithms: TargetScan v7.1<sup>2</sup>, miRanda<sup>3</sup>, DIANA-TarBase v7.0<sup>4</sup>, and PicTar<sup>5</sup> via the "miRNAatp" package (Dhawan et al., 2018).

## Cell Culture and Transfection of miR-200a-3p Mimic and Inhibitor

Human colon cancer cell lines DLD1 and SW480 and normal intestinal epithelium cell line NCM460 were provided by the Cell Bank of Wuhan University (Wuhan, China). hsa-miR-200a-3p mimic, inhibitor, mimic NC, and inhibitor-negative control were from RiboBio Co., Ltd. (Guangzhou, China). Cell culture and transient transfection were performed as described previously (Sun et al., 2017a; Lin et al., 2019).

## Tissue Collection

Colorectal cancer tissues and ANT (distance to cancer >5 cm) were sampled from sixty patients with primary CRC that was diagnosed by pathological assessment of tissues at Taihe Hospital of Hubei University of Medicine from January 2017 to December 2018. The study was authorized by the Research Ethics Committee of Hubei University of Medicine (Shiyan, Hubei, China). Informed consent was provided by all patients who participated in the present study.

## Total RNA Isolation and Quantitative Reverse Transcription Polymerase Chain Reaction (qRT-PCR)

Total RNA was isolated using Trizol reagent (Invitrogen, United States). The expression of mature miR-200a-3p was determined using qRT-PCR with the Bulge-Loop<sup>TM</sup> miRNA qRT-PCR Primer Set and Control Primer Set (RiboBio, Guangzhou, China); 2 μg of RNA was used to synthesize cDNA using the RevertAid First Strand cDNA Synthesis Kit (Thermo, United States). qRT-PCR was performed using the FastStart Universal SYBR Green Master (Rox) (Roche, United States) in the ABI PRISM<sup>®</sup> 7300 real-time PCR system (Applied Biosystems, Foster City, CA, United States). GAPDH and U6 were added as endogenous controls, and melting curves were used to check non-specific amplifications. Relative expression level was calculated by the  $2^{-\Delta\Delta C_t}$  method. The primer sequences are shown in **Supplementary Data Sheet S1**.

## Western Blotting and Dual-Luciferase Reporter Assays, Colony Formation Assay, Wound Healing Assay *in vitro*, and Transwell Migration/Invasion Assay

Approximately 0.2 g CRC and paired normal tissue samples were ground in liquid nitrogen and lysed with RIPA lysis buffer (CellLytic, Sigma-Aldrich, United States) and proteinase inhibitor cocktail (Merck Millipore, United States). Total protein was

<sup>2</sup>[http://www.targetscan.org/vert\\_71/](http://www.targetscan.org/vert_71/)

<sup>3</sup><http://www.microrna.org/microrna/home.do>

<sup>4</sup><http://diana.imis.athena-innovation.gr/DianaTools/index.php?r=tarbase/index/>

<sup>5</sup>[http://pictar.mdc-berlin.de/cgi-bin/PicTar\\_vertebrate.cgi](http://pictar.mdc-berlin.de/cgi-bin/PicTar_vertebrate.cgi)

extracted and immunoblotted as described previously (Sun et al., 2017a; Lin et al., 2019). The antibody used in the present study is shown in **Supplementary Data Sheet S1**. The dual-luciferase reporter assays, colony formation assay, wound healing assay *in vitro*, transwell migration/invasion assay, transfection reagents, primers, and western blot antibody are described in **Supplementary Data Sheet S1**.

## Statistical Analysis

All statistical analyses were performed using R/BioConductor (version 3.5.1) with two-tailed *P*-values. In the training phase, the dependent variable of the LASSO-logistic regression and SVM-RFE algorithm is dichotomized into two groups (tumor vs. normal), whereas the independent variable is the continuous expression values of 68 miRNAs.

In the testing phase, logistic regression is used for the diagnostic model of the binary variables, tumor and normal, and the independent variable is the continuous expression values of nine miRNAs. ROC was analyzed with R (“pROC,” “MASS,” “nnet,” “Daim,” and “epicalc” packages) to evaluate the diagnostic efficiency for the diagnostic model on the CRC outcome for the expression level of the nine candidate miRNAs. For Cox regression, the dependent variable is survival data showing the state of survival and the time of follow-up, whereas the independent variable is the continuous expression values of the nine miRNAs. Kaplan-Meier plots and ROC curve for the prognosis model were constructed using R (“survival,” “survminer,” “ggplot2,” and “survivalROC” packages). The median value was used as the cutoff value between low and high expression levels of each miRNA. The Cox proportional hazards regression model was applied to determine OS, and log-rank test was used to detect differences in OS.

For the validation phase, the nomogram prognostic model contains clinical features such as the total TNM Stage, T stage, N stage, M stage, and risk score. The clinical features (TNM Stage, T stage, N stage, M stage) are factor variables (categorical variables), and risk score is a continuous variable. The performance of the nomogram for the testing and validation data was evaluated by the concordance index and was distributed graphically with calibration plots using the “rms” package in R.

In the verification phase, ANOVA or Student’s *t*-test was used to detect significant differences between the groups with *P* < 0.05 as the threshold. The data from at least three independent experiments were averaged and presented as mean ± SD.

## RESULTS

### Identification of Significantly Deregulated miRNAs

Three CRC miRNA datasets consisting of 170 CRC samples and 172 ANT samples (**Supplementary Table S1**) were selected by integrated analysis (**Figure 1**), from which 135 miRNAs with steady DE patterns were identified by a moderated *t*-test with addition of a fudging parameter, and Fisher’s method with summarization of  $-\log(p\text{-value})$  across studies to run 300 selections (Wang et al., 2019). The effect size was merged,

and a total of 68 DE miRNAs with profiles of DE miRNAs similar to those from Fisher’s method of combining *P* < 0.001 (**Supplementary Table S2**) were determined (**Figure 2A**). As expected, the CRC was differentiated from ANT samples by hierarchical clustering of the 68 DE miRNAs (**Figure 2B**).

### Screening of Candidate miRNAs

The miRNAs that were selected by the LASSO (**Figure 3A**) and SVM-RFE (**Figure 3B**) algorithms were merged, among which nine miRNAs that were overlapped in both algorithms (**Figure 3C**, **Supplementary Table S3** and **Supplementary Figure S1**) and constituted a minimum candidate list for model prediction of CRC samples (**Figures 3A–C**).

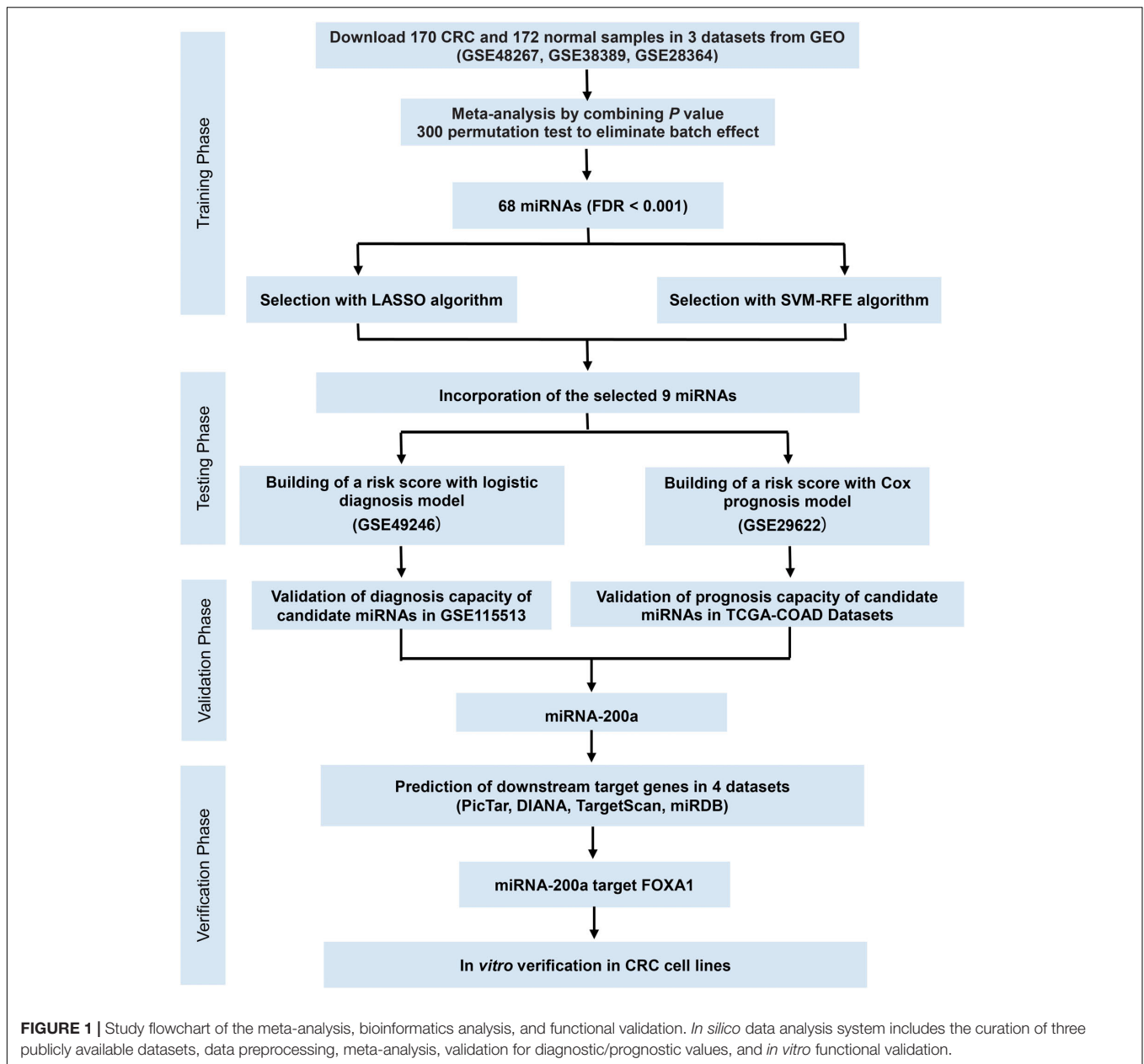
### Diagnostic Capacity of Candidate miRNAs for CRC in GEO Datasets

A total of 80 testing-phase samples in GSE49246 and 1513 validation-phase samples in GSE115513 were differentiated into the CRC and normal groups using a linear logistic regression model of nine miRNAs. A DRS was estimated and weighted with the coefficients by the univariate linear regression model to evaluate the performance of nine miRNAs for prediction of CRC in the GSE49246 cohort. The DRS was estimated by the formula:  $DRS = (-5.1928 \times EXP_{hsa-miR-492}) + (0.1885 \times EXP_{hsa-miR-200a}) + (-3.5434 \times EXP_{hsa-miR-338}) + (-9.1146 \times EXP_{hsa-miR-29c}) + (0.8177 \times EXP_{hsa-miR-101}) + (-1.3238 \times EXP_{hsa-miR-148a}) + (2.9460 \times EXP_{hsa-miR-92a}) + (-1.2699 \times EXP_{hsa-miR-424}) + (2.7264 \times EXP_{hsa-miR-210})$ . ROC curve analysis indicated high categorization accuracy of a panel of nine miRNA signatures in the testing set GSE49246 (AUC = 0.94) (**Figure 3D**), in the validation set GSE115513 (AUC = 0.89) (**Figure 3E**), and in our independent validation set (AUC = 0.978) (**Figure 3F** and **Supplementary Table S4**).

### Prognostic Capacity of Candidate miRNAs for CRC in GEO and TCGA Datasets

Significant associations of candidate miRNAs with the survival of CRC patients were analyzed by multivariate Cox regression in 80 testing-phase GSE29622 samples (**Supplementary Table S4**). PRS was estimated using the nine-miRNA signature model as follows:  $PRS = (0.1170 \times EXP_{hsa-miR-492}) + (0.3496 \times EXP_{hsa-miR-92a}) + (0.3640 \times EXP_{hsa-miR-424}) + (-0.5316 \times EXP_{hsa-miR-29c}) + (0.0573 \times EXP_{hsa-miR-101}) + (0.0882 \times EXP_{hsa-miR-338}) + (-0.5038 \times EXP_{hsa-miR-200a}) + (0.1920 \times EXP_{hsa-miR-210}) + (0.2397 \times EXP_{hsa-miR-148a})$ , where EXP was the FPKM value of the miRNA.

The prognostic performance of the nine-miRNA signature model was characterized using the GSE29622 dataset. The median PRS (1.036) was used to classify CRC patients into a low-risk group (*n* = 33) and a high-risk group (*n* = 32) (**Figure 4A**). As shown in **Figure 4B**, the high-risk group had significantly shorter survival time or lower survival probability compared with the low-risk group (HR, 4.741, 95% CI, 2.13–10.55, log-rank test *p* = 0.0002). The AUC for predicting the 1- to 10- year survival of CRC patients was between 0.783 and 0.872 (**Figure 4C**). High



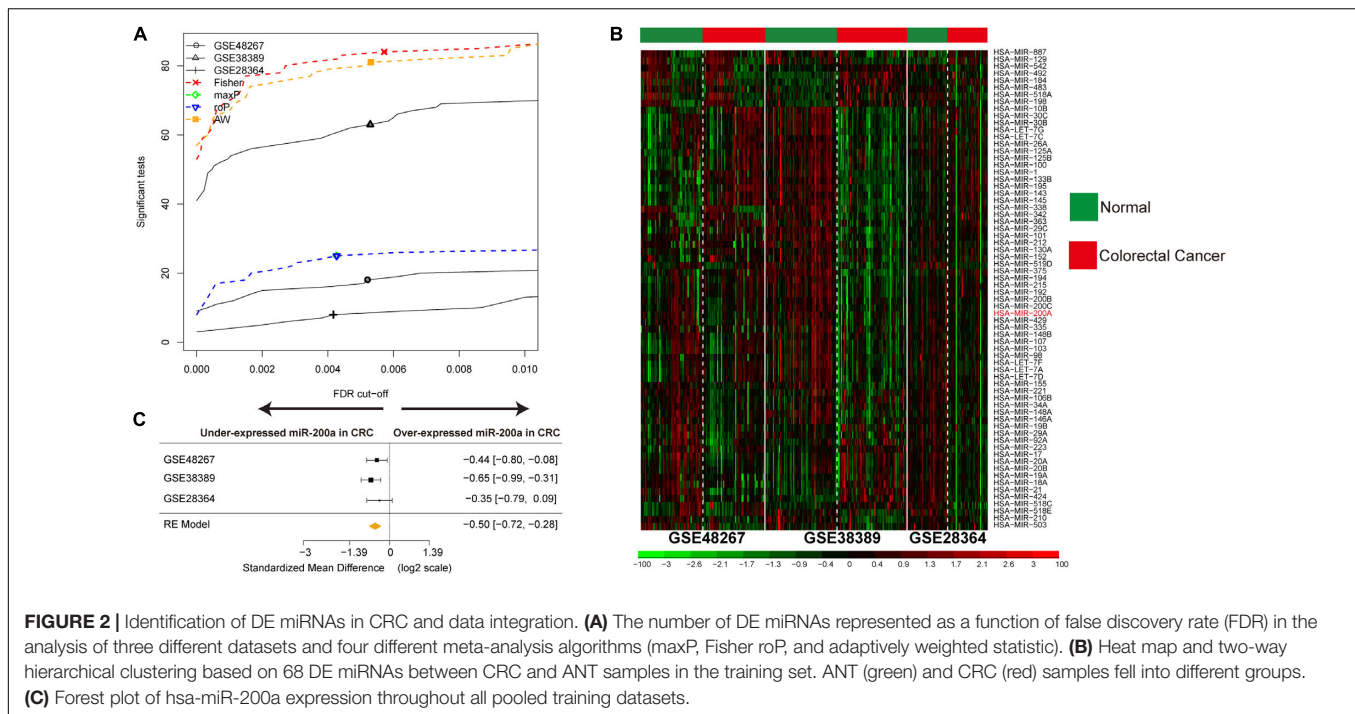
expression of hsa-miR-200a and low expression of hsa-miR-492, hsa-miR-92a, hsa-miR-424, hsa-miR-29c, hsa-miR-101, hsa-miR-338, hsa-miR-210, and hsa-miR-148a were associated with longer predicted survival (Figures 4D–L). The results suggested the nine-miRNA signature enabled effective prediction of the prognosis of CRC patients.

### Validation of Prognostic Performance of the Nine-miRNA Signature in TCGA-COAD

The prognostic performance of the nine-miRNA prediction model for CRC was validated using the validation dataset TCGA-COAD. The CRC patients were classified into the high-risk group

( $n = 203$ ) and the low-risk group ( $n = 203$ ) using PRS and the median PRS as cut-off criteria (Figure 5A). The high-risk group had significantly shorter survival time or lower survival probability compared with the low-risk group (HR, 3.9, 95% CI, 2.125–7.16, log-rank test  $p < 0.001$ ) (Figure 5B). The AUC for predicting 1- to 10- year survival of CRC was between 0.796 and 0.911 (Figure 5C). The results were consistent with those of the testing set (Figures 4, 5), supporting the effectiveness of the nine-miRNA signature in predicting the prognosis of CRC.

The effect of each clinicopathological feature on survival was analyzed by Cox regression to evaluate whether the prognostic classifier acted as an independent indicator in CRC patients. As shown in Table 1, after multivariable adjustments for clinicopathological factors, the nine-miRNA-based classifier was



demonstrated to be a powerful and independent factor in the testing cohort of 65 cases (HR 1.232, 95% CI 1.115–1.363,  $P < 0.001$ ) and in the validation set of 406 cases (HR 1.455, 95% CI 1.287–1.645;  $P < 0.001$ ).

Time-dependent ROC curve analysis revealed that the nine-miRNA signature was a better predictor of survival compared with clinicopathological factors in both the testing and validation sets (Figures 5D,E).

## Building a Predictive Nomogram and Calibration Plot

A predictive model applicable for clinics was built using a nomogram to predict an individual's 1-, 3-, and 5-year survival, taking into consideration clinicopathological covariates. According to multivariate analysis of OS (Table 1), a nomogram with predictors including PRC, TMN stage, and pathological stage was used to predict the 1-, 3-, 5-, and 7-year survival in both the testing cohort (Supplementary Figure S2A) and validation cohort (Supplementary Figure S2B). The concordance index was 0.827 (0.761–0.893) for the testing cohort (Supplementary Figure S2C) and 0.837 (0.778–0.895) for the validation cohort (Supplementary Figure S2D), suggesting that the calibration plots effectively predicted the 1-, 3-, and 5-year OS rate.

## Association of Downregulated miR-200a-3p With Upregulated FOXA1 in CRC

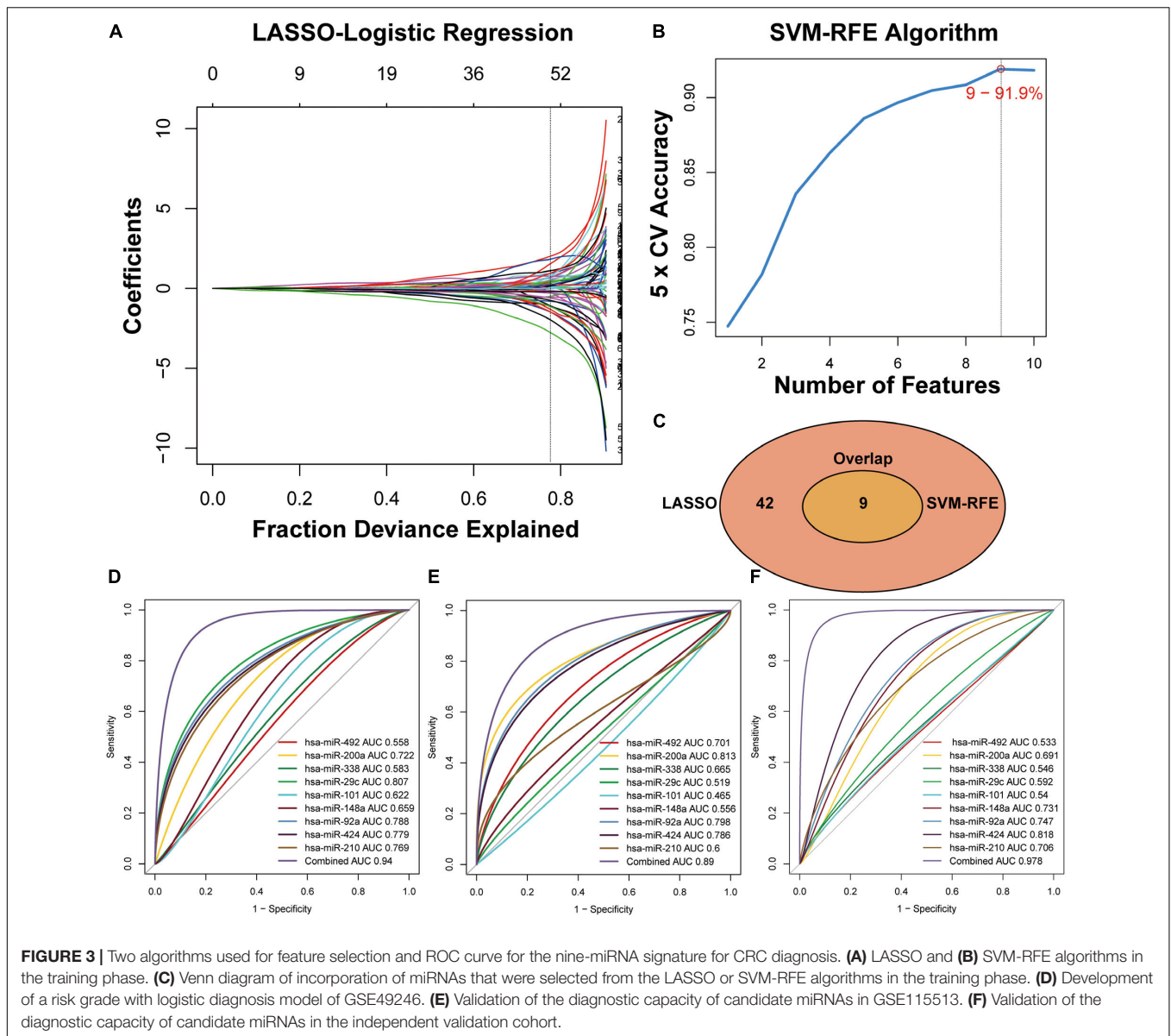
A decreasing mode in the miR-200a levels was observed in human CRC (Figure 2C). The levels of miR-200a-3p in DLD1

and SW480 cell lines were lower than those in the normal intestinal epithelium cell line NCM460 (Figure 6A). qRT-PCR (normalized with U6) was used to measure mature miR-200a-3p level in the independent validation sample cohort ( $n = 60$ ). The results indicated that the expression of miR-200a-3p was significantly ( $P < 0.05$ ) decreased in the ANT-related tumors (Figure 6B). Both qRT-PCR and Western blot analyses demonstrated significant upregulation of FOXA1 in CRC tissues compared with ANT (Figures 6C,D). The miR-200a-3p expression levels were significantly and inversely associated with FOXA1 levels in both CRC tissues and normal samples ( $R^2 = 0.21$ ,  $P = 1.01E-07$ ) (Figure 6E) in our independent validation cohort, but they were not associated with FOXA1 levels in the TCGA colon or rectal cancer cohort (Supplementary Figure S3).

A total of 66 common targets of miR-200a were identified from both prediction algorithms and experimentally supported databases using four publicly available algorithms with high-stringency (Figure 6F and Supplementary Table S5).

## Inhibition of miR-200a in Clone Formation, Migration, and Invasion of Colon Cancer Cells

Based on the results of the integrated analysis and *in vitro* experiments, we hypothesized that the reduced expression of miR-200a-3p might accelerate CRC tumorigenesis via targeting FOXA1. We transfected SW480 and DLD1 cells with miR-200a-3p mimic, miR mimic NC, miR-200a-3p inhibitor, miR inhibitor NC to determine the role of miR-200a-3p in clone formation. Clonogenic assay revealed that the clonogenic survivals of SW480 and DLD1 cells were reduced by miR-200a-3p compared with the NC groups, whereas phenotype

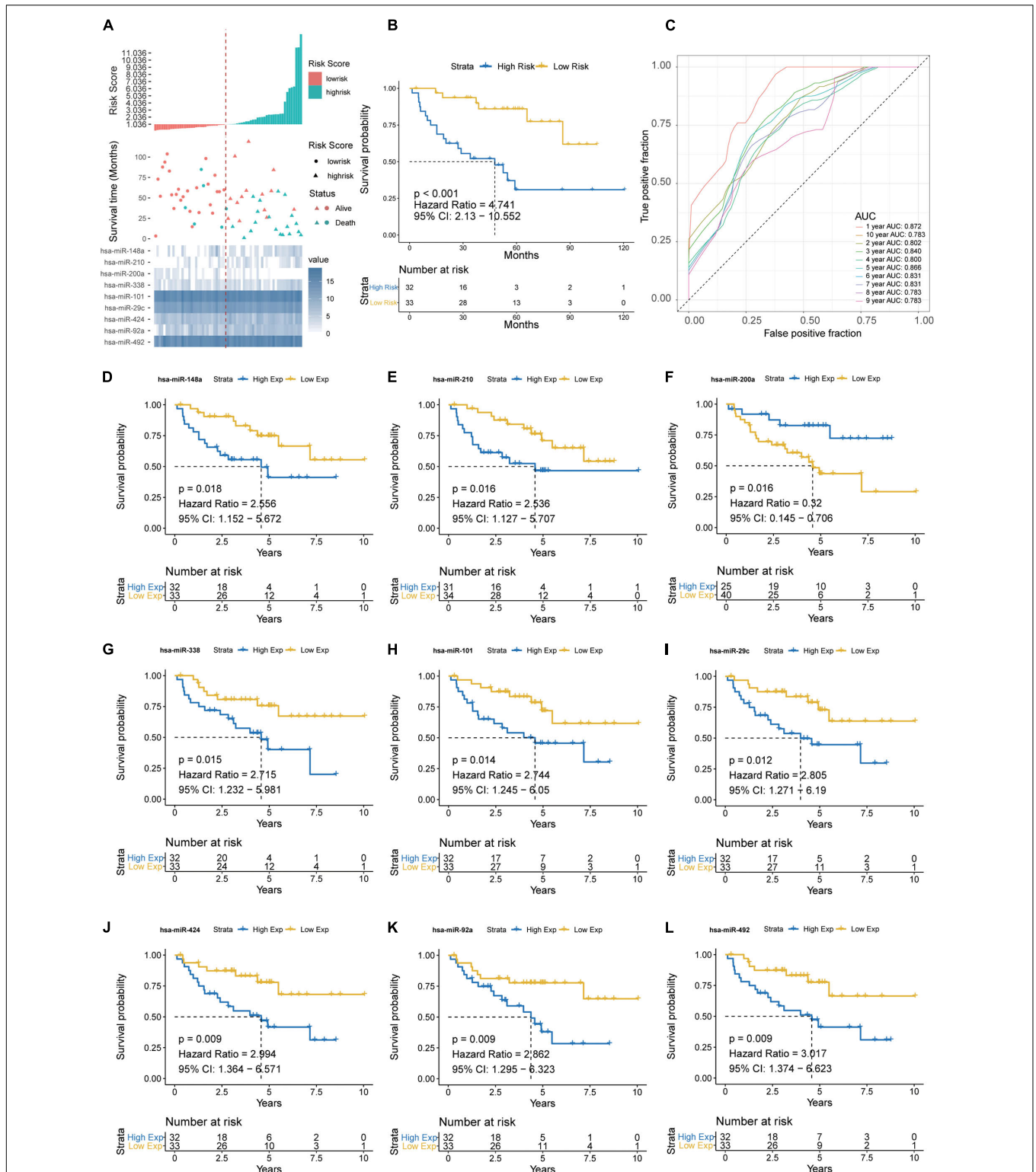


was reversed by miR-200a-3p inhibitor (Figures 7A,B). The wound healing assay showed that miR-200a-3p overexpression suppressed and miR-200a-3p inhibitor enhanced migration of CRC cells (Figures 7C–F). The transwell assay showed that miR-200a-3p overexpression suppressed and miR-200a-3p inhibitor enhanced invasion of CRC cells (Figures 7G–J). These data indicated that miR-200a suppressed the proliferation, migration, and invasion of CRC cells.

### miR-200a-3p Modulates FOXA1 and YAP1 Expression via Targeting Human FOXA1 3'UTR and Inhibits EMT of Colon Cancer Cells

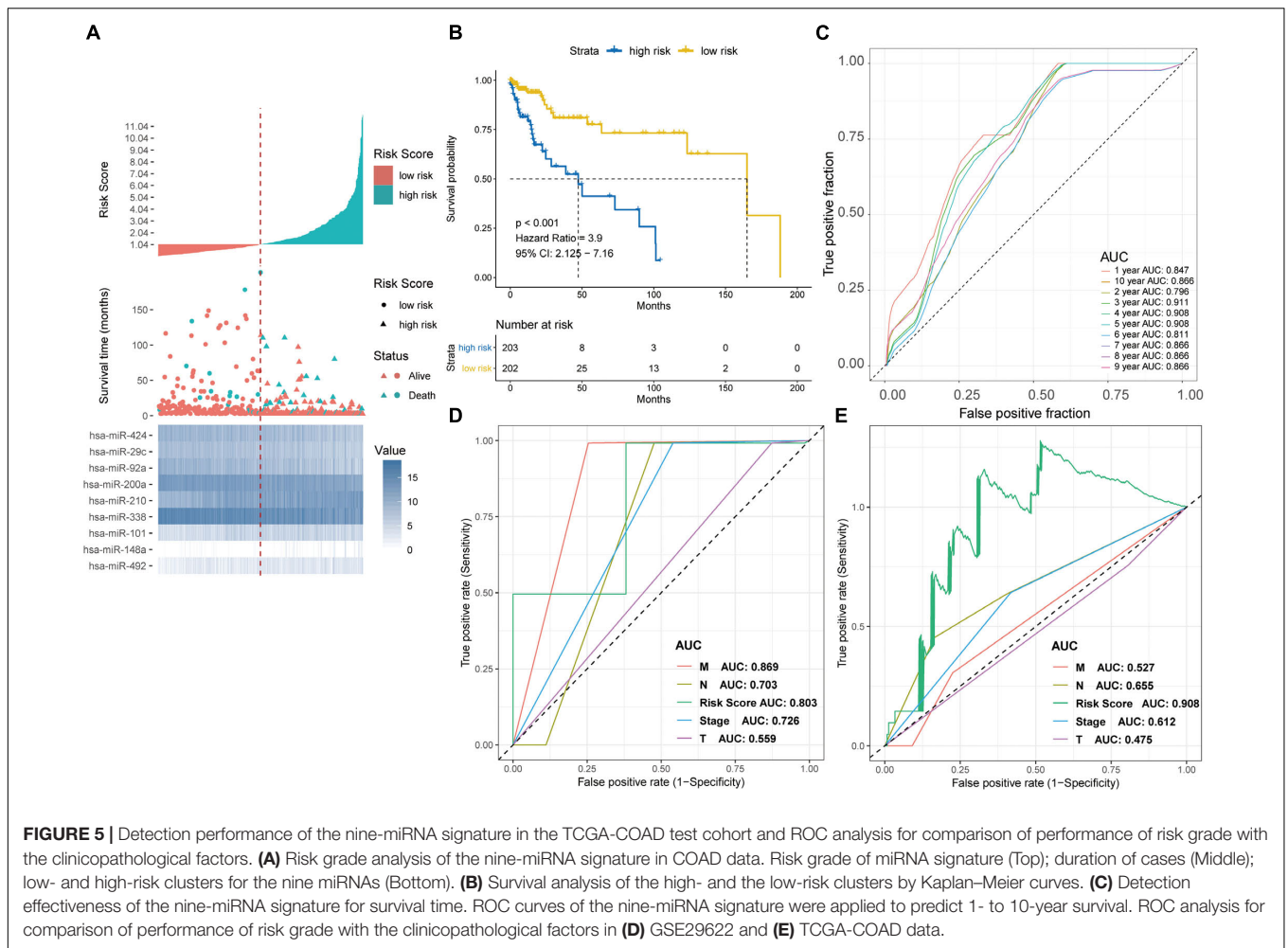
FOXA1 3'-UTR was cloned into a luciferase reporter plasmid (Figure 8A), and expression of the adjacent hLuc coding

region was quantified to evaluate miR-200a-3p-regulated FOXA1 expression. Searching different databases revealed that FOXA1 is a predicted target of miR-200a-3p that presents oncogenic properties. FOXA1 harbors two conserved miR-200a-3p cognate sites: nucleotides 310–317 and 271–276 of FOXA1 3'-UTR (Figures 8B,C). The luciferase reporter plasmid FOXA1-3'-UTR or mutant reporter plasmids with point mutations in the putative miR-200a-3p binding sites were co-transfected with miR-200a-3p mimics or miR mimic NC and inhibitors, separately. The results showed that the luciferase activity of reporter plasmid with wild type FOXA1 3'-UTR was significantly inhibited by miR-200a-3p but increased by miR-200a-3p inhibitor (Figure 8D,  $P < 0.05$ ). However, no significant change was observed in reporter plasmid containing mutated FOXA1 3'-UTR (i.e., MUT-FOXA1-3'-UTR). These results suggested that miR-200a-3p negatively regulated FOXA1



**FIGURE 4 |** Detection performance of the nine-miRNA signature in GSE29622 data. **(A)** Risk grade analysis of the nine-miRNA signature of CRC. Risk grade of miRNA signature (Top); duration of cases (Middle); low- and high-risk clusters for the nine miRNAs (Bottom). **(B)** Survival analysis of the high-risk cluster and the low-risk cluster with Kaplan–Meier curves. **(C)** Detection effectiveness of the nine-miRNA signature for survival time. ROC curves of the nine-miRNA signature were used to predict 1- to 10-year survival. Independent detection power (efficiency) of individual miRNAs in the nine-miRNA signature. **(D)** hsa-miR-148a, **(E)** hsa-miR-210, **(F)** hsa-miR-200a, **(G)** hsa-miR-338c, **(H)** hsa-miR-101, **(I)** hsa-miR-29c, **(J)** hsa-miR-424, **(K)** hsa-miR-92a, and **(L)** hsa-miR-492. The x-axis represents overall survival time, whereas the y-axis indicates OS probabilities.



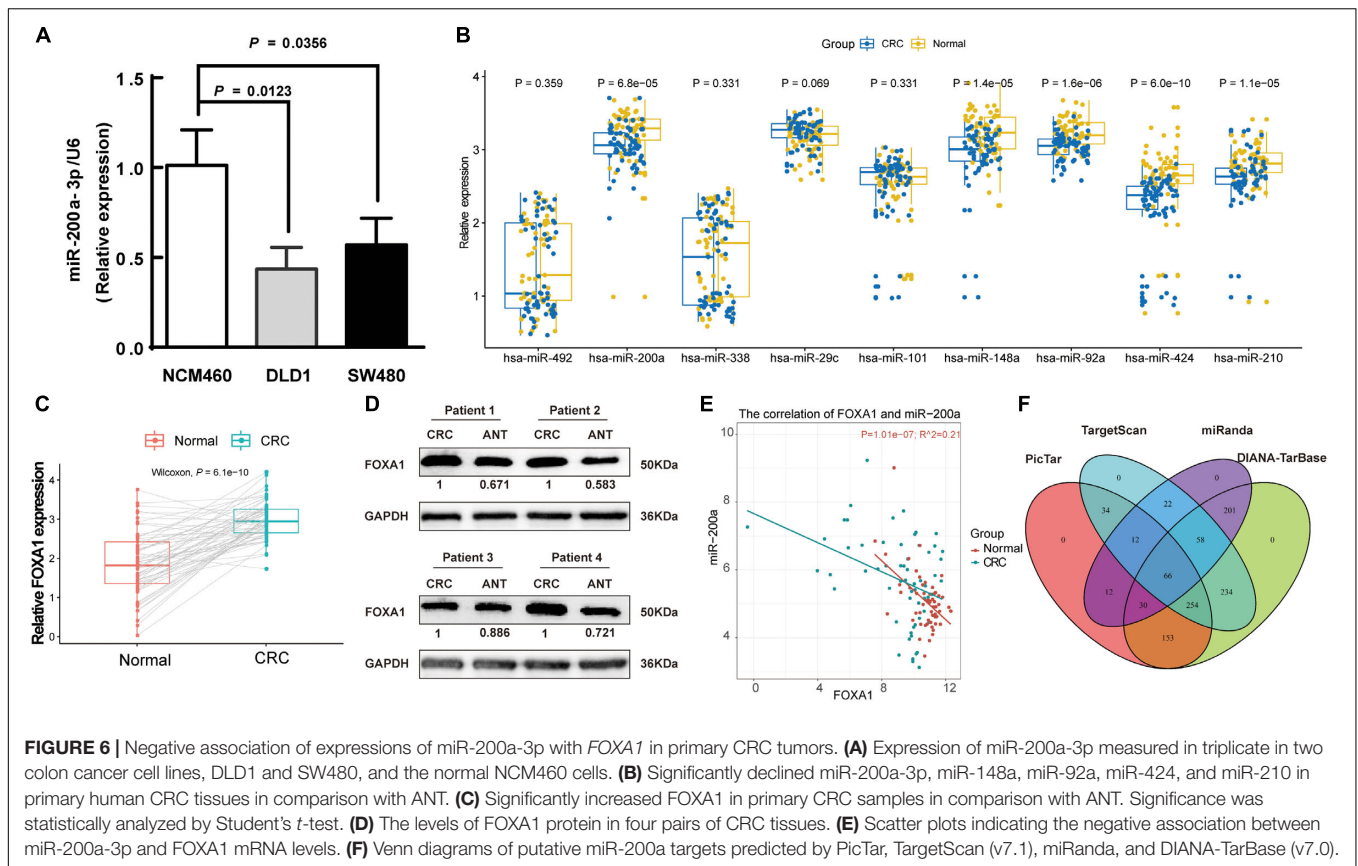


**FIGURE 5 |** Detection performance of the nine-miRNA signature in the TCGA-COAD test cohort and ROC analysis for comparison of performance of risk grade with the clinicopathological factors. **(A)** Risk grade analysis of the nine-miRNA signature in COAD data. Risk grade of miRNA signature (Top); duration of cases (Middle); low- and high-risk clusters for the nine miRNAs (Bottom). **(B)** Survival analysis of the high- and the low-risk clusters by Kaplan–Meier curves. **(C)** Detection effectiveness of the nine-miRNA signature for survival time. ROC curves of the nine-miRNA signature were applied to predict 1- to 10-year survival. ROC analysis for comparison of performance of risk grade with the clinicopathological factors in **(D)** GSE29622 and **(E)** TCGA-COAD data.

**TABLE 1 |** Univariate and multivariate analyses of clinicopathological factors and nine-miRNA classifier with OS in GSE29622 and TCGA COAD cohorts.

	Univariate analysis <sup>a</sup>		Multivariate analysis <sup>b</sup>	
	HR (95% CI)	P-value	HR (95% CI)	P-value
<b>GSE29622 testing set (n = 65)</b>				
Gender (male vs. female)	1.586 (0.682–3.687)	0.284		
Pathologic T stage (T3–4 vs. T1–2)	1.678 (0.394–7.141)	0.484	0.559 (0.099–3.162)	0.511
Pathologic N stage (N0 vs. N1 vs. N2)	<b>2.744 (1.562–4.820)</b>	<b>&lt;0.001</b>	1.322 (0.559–3.128)	0.525
Pathologic M stage (present vs. absent)	<b>5.059 (2.243–11.410)</b>	<b>&lt;0.001</b>	<b>3.09 (1.231–7.757)</b>	<b>0.016</b>
TNM stage (III and IV vs. I and II)	<b>5.290 (1.813–15.436)</b>	<b>0.002</b>	2.331 (0.421–12.911)	0.332
Risk Score (continuous value)	<b>1.272 (1.159–1.396)</b>	<b>&lt;0.001</b>	<b>1.232 (1.115–1.363)</b>	<b>&lt;0.001</b>
<b>TCGA-COAD validation set (n = 406)</b>				
Gender (male vs. female)	1.299 (0.727–2.321)	0.376		
Pathologic T stage (T3–4 vs. T1–2)	<b>3.081 (1.393–6.818)</b>	<b>0.005</b>	2.188 (0.919–5.210)	0.077
Pathologic N stage (N0 vs. N1 vs. N2)	<b>1.541 (1.092–2.175)</b>	<b>0.014</b>	1.654 (0.897–3.050)	0.107
Pathologic M stage (present vs. absent)	<b>3.919 (2.080–7.383)</b>	<b>&lt;0.001</b>	<b>3.335 (1.511–7.358)</b>	<b>0.003</b>
TNM stage (III and IV vs. I and II)	<b>1.839 (1.027–3.292)</b>	<b>0.040</b>	0.602 (0.189–1.916)	0.390
Risk score (continuous value)	<b>1.415 (1.262–1.587)</b>	<b>&lt;0.001</b>	<b>1.455 (1.287–1.645)</b>	<b>&lt;0.001</b>

<sup>a</sup>The data were subjected to the Cox proportional hazards regression model. <sup>b</sup>Multivariate analysis used stepwise addition and removal of clinical covariates found to be associated with survival in univariate models ( $P < 0.05$ ), and final models include only those covariates that were significantly associated with survival (Wald statistic,  $P < 0.05$ ). Bold indicate statistically significant values ( $P < 0.05$ ).



expression by directly binding to the predicted binding site(s) in the *FOXA1* 3'-UTR.

DLD1 and SW480 cells were transfected with miR-200a-3p mimic, miR mimic NC, miR-200a-3p inhibitor, and inhibitor NC to verify the direct regulation of *FOXA1* expression by miR-200a-3p. Both qRT-PCR and Western blotting revealed that the expression level of *FOXA1* was suppressed in miR-200a-3p-overexpressed cells but regained in miR-200a-3p inhibitor-treated cells (Figures 8E,F). Therefore, miR-200a-3p negatively regulated *FOXA1* levels in CRC cells. In addition, the results showed that miR-200a-3p negatively regulated expression of N-cadherin and YAP1 proteins and positively regulated E-cadherin expression (Figure 8F), suggesting that miR-200a-3p inhibited EMT of CRC cells.

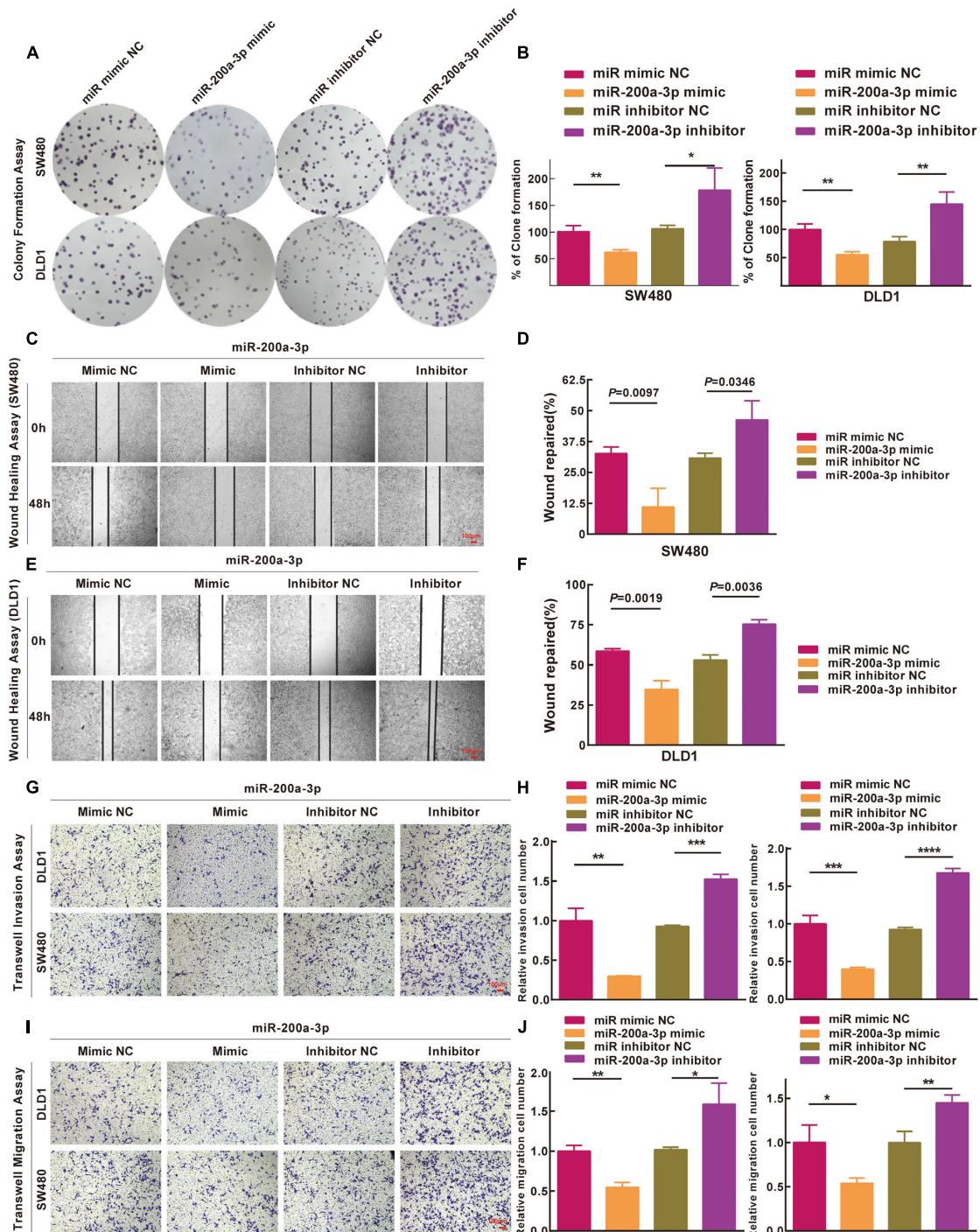
## DISCUSSION

Integrated analysis that unites cross-platform normalization of gene profiling and neglects the unavailability of raw data has been widely applied to detect DEGs at the mRNA and miRNA levels in CRC (Sun et al., 2017a,b; Lin et al., 2019). In the present study, the microarray data were analyzed by meta-analysis to detect candidate miRNAs that were potential diagnostic and prognostic markers. Further, a novel diagnostic and prognostic model based on candidate miRNAs was developed and validated for OS prediction of CRC in both testing and validation cohorts.

The novel tool successfully differentiated patients into high-risk and low-risk clusters with significant differences in OS. The classifier was proven to predict the survival of patients with CRC significantly better than other clinicopathological risk factors. When stratified by these, the nine-miRNA-based classifier appears to be an effective prognostic model to provide value supplementary to clinicopathological characteristics. The nine-miRNA signature was also demonstrated to be independent of other clinicopathological factors. In addition, the prognostic scores can be easily calculated according to the prognostic-score equation, and the levels of the nine miRNAs in the signature can be easily determined using quantitative real-time PCR or other customized test methods. Risks and prognosis can be effectively predicted by the prognostic scores. Therefore, the prognostic score model is easily applied by oncologists in clinical practice.

A comprehensive and practical nomogram was developed to predict individual survival rates, providing opportunities for clinicians to rank patients according to these valid tools. miR-200a-3p is the only tumor-suppressor miRNA of the nine-miRNA-based classifier and inhibits expression of *FOXA1*, *YAP1*, and N-Cadherin but enhances E-Cadherin expression in colon cancer cells. The expression of *FOXA1* was directly regulated by miR-200a-3p binding to the 3'-UTR of *FOXA1*.

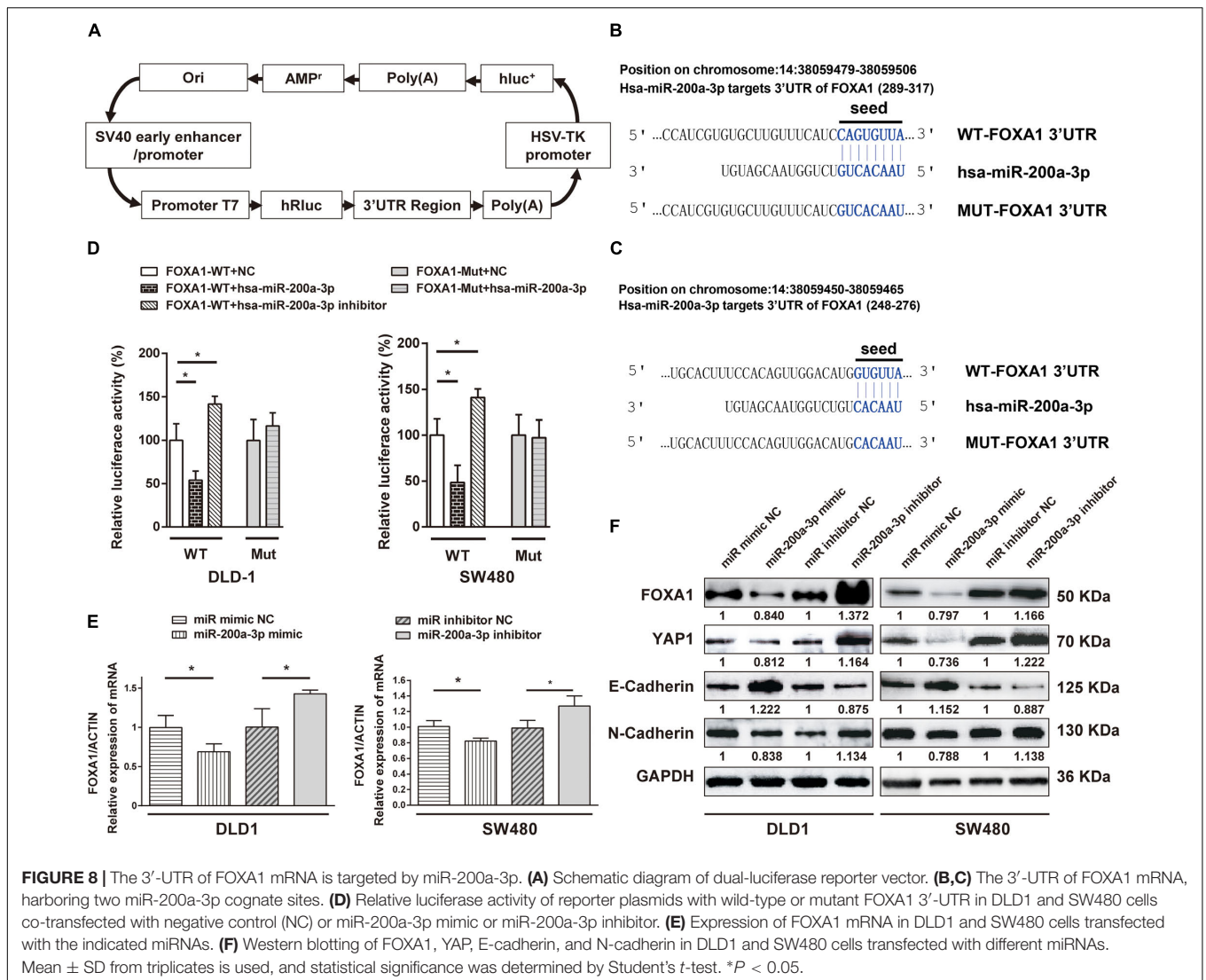
Zhang et al. (2013) used a tool-integrated LASSO Cox regression model with multiple miRNAs to develop a molecular classifier based on a core set of six miRNAs for predicting the 5-year disease-free survival for patients with stage II CRC.



**FIGURE 7 |** Inhibitions of miR-200a-3p to formation, movement, and penetration in SW480 and DLD1 cell clone *in vitro*. **(A,B)** Typical photomicrographs and quantifications of clone formation assay in SW480 and DLD1 cells at 48 h after transfection. **(C–F)** Photomicrographs of wound healing assay in SW480 and DLD1 cells at 48 h after transfection. **(G)** Transwell invasion assays of DLD1 and SW480 cells. Bar = 100  $\mu$ m. **(H)** Total number of cells in five fields was estimated manually. **(I)** Transwell migration assays of DLD1 and SW480 cells with different miRNAs. Bar = 100  $\mu$ m. **(J)** Total number of cells in five fields was determined manually. Mean  $\pm$  SD was used, and statistical significance analysis was performed with one-way ANOVA. \* $P < 0.05$ . \*\* $P < 0.01$ . \*\*\* $P < 0.001$ .

However, previous studies on miRNA diagnostic and prognostic prediction were restricted by the small number of miRNAs screened, small sample sizes, lack of independent validation, and

utilization of improper statistical methods for miRNA microarray data mining. In the present study, all nine candidate miRNAs were detected by SVM-RFE instead of LASSO; this difference



might be caused by the inherent characteristics of each set of algorithms. It should be noted, however, that the nine miRNAs selected by integrated functions of LASSO and SVM-RFE were further validated, suggesting the feasibility of the integration strategy. Our model still needs to be validated with real samples in different oncology centers. In addition, all nine miRNAs except miR-101 were newly detected to be associated with CRC. Further characterization of molecules such as miR-210 will provide novel insights into CRC occurrence and advancement, leading to the identification of potential therapeutic targets for patients with CRC.

A number of miRNAs have been detected to predict the survival of patients with CRC (Balaguer et al., 2011; Zhang et al., 2013; Hur et al., 2015; Vychytilova-Faltejskova et al., 2016; Pettit et al., 2017; Toiyama et al., 2017; Ji et al., 2018; Kandimalla et al., 2018). Particularly, miR-492, miR-200a, miR-338, miR-29c, miR-101, miR-148a, and miR-92a have proven to be correlated with the prognosis or therapeutic outcome of patients with CRC (Gaedcke et al., 2012; Kuo et al., 2012;

Schee et al., 2012; Tsai et al., 2013; Yong et al., 2013; Kim et al., 2014; Pichler et al., 2014; Sun et al., 2014; Xue et al., 2014; Chen et al., 2018; Fu et al., 2018; Igder et al., 2019). miR-492 is a rectal cancer-specific miRNA that was the single most upregulated (up to 16-fold higher) in rectal cancer and was never detected for colon cancers (Gaedcke et al., 2012). In an effort to characterize miRNAs related to aggressiveness and E-cadherin expression, Pichler et al. (2014) revealed low expression level of miR-200a as an independent prognostic factor pertaining to cancer-specific survival, contrary to the expression of EMT-related genes in cancer stem cell-enriched spheroid growth conditions. Yong et al. (2013) proposed that a three-miRNA (miR-193a, miR-23a, and miR-338) classifier appeared to be a potential blood biomarker for early detection of CRC. Xue et al. (2014) reported that miR-338-3p inhibited colorectal carcinoma cell invasion and migration by attacking smoothened. Kuo et al. (2012) reported significant downregulation of miR-29a and miR-29c in early-recurring CRC patients and proposed miR-29a/c as promising biomarkers for CRC early recurrence.

Yang et al. (2019) demonstrated that miR-101 inhibited colon cancer by targeting CREB1. Tsai et al. (2013) discovered that hsa-miR-148a inhibited tumorigenesis by suppressing colon cancer cell multiplication and migration but not invasion, resulting in an accumulation of the G2 population. Therefore, hsa-miR-148a might be a simple and accountable biomarker for the detection of postoperative early relapse in patients with CRC after radical resection (Tsai et al., 2013). Overexpression of miR-92a in HCT116 cells was demonstrated to facilitate migration, proliferation, and resistance to apoptosis and to modify F-actin organization, leading to the enhanced oncogenicity (Alcantara and Garcia, 2019).

Chen et al. (2012) demonstrated that miR-424 overexpression significantly abrogated invasive potential, acting as a direct tumor suppressor. In the present study, we found that only a high expression level of miR-200a-3p in tumors was significantly associated with the enhanced OS of CRC patients, indicating a negative correlation of miR-200a-3p to CRC progression. Pichler et al. (2014) found inverse expression of EMT-related genes caused by abnormal expression of miR-200a, but the study did not explore its potential mechanism in CRC.

Considering the convenience of clinical application, miR-200a-3p was selected for function validation in our study. We confirmed downregulation of miR-200a in CRC tissues, which was substantially linked to poor survival of CRC patients. miRNA target analysis and luciferase reporter assays suggested that the miR-200a-3p-mediated regulatory mechanism in colon cancer cells was inhibition of *FOXA1* expression by targeting the 3'-UTR of *FOXA1* mRNA. Over-expression of miR-200a-3p significantly suppressed cellular proliferation. Furthermore, ectopic overexpression of miR-200a-3p reduced cell migration and invasion and enhanced expression of E-cadherin, suggesting the potential role of miR-200a-3p in CRC metastasis. To the best of our knowledge, this is the first meta-analysis to reveal the comprehensive mechanism underlying malignant development of CRC caused by loss of miR-200a-3p, which unleashes expression of *FOXA1*. *FOXA1* is a transcription factor. *FOXA1* knockdown distinctly suppressed cell proliferation and promoted apoptosis in SW480 and HCT116 CRC cells. The expression of YAP in SW480 and HCT116 cells was also prominently downregulated by *FOXA1* knockdown to reduce the tumor growth of CRC (Ma et al., 2016), consistent with our previous study that confirmed inhibition of down-expression of YAP1 to the EMT in a manner associated with CRC invasion *in vitro* (Sun et al., 2017a). The Hippo/YAP signature was greatly enriched for *FOXA1* binding sites, ranking second among all transcription factors analyzed by Fitamant et al. (2015). Therefore, the interaction between *FOXA1* and YAP1 represents the balance between tumor maintenance and cell discrimination (Fitamant et al., 2015). Further study is needed to reveal more functions of miR-200a-3p in CRC carcinogenesis and development by targeting the *FOXA1/YAP1* signal pathway.

In the current study, we found that expression of miR-200a-3p was significantly decreased in ANT-related tumors and a significant upregulation of *FOXA1* in CRC tissues compared with ANT. The miR-200a-3p expression levels were significantly

and inversely associated with *FOXA1* levels in both CRC tissues and normal samples in our independent validation cohort, but they were not associated with *FOXA1* levels in the TCGA colon or rectal cancer cohort. Different ethnic groups, genetic testing method, dietary habits, and culture may be responsible for the inconsistencies between the TCGA datasets and our independent validation cohort. Firstly, in terms of ethnic group, the original data source of the independent validation cohort was Asian, whereas the data source for TCGA datasets is Caucasian. Second, the data are detected in different ways. The assay used in the independent validation cohort is qRT-PCR, and the unit is the ratio of the internal parameters. The data of the TCGA datasets is sequencing data, and the unit of measurement is RPKM (Reads Per Kilobase of exon Model per Million mapped reads). Third, CRCs from patients embedded in geographically diverse populations and cultures reflect substantially different dietary exposures, extended over the whole life course. These differences may account for the inconsistency between the expression of miR-200a and *FOXA1* in TCGA and that of the independent validation cohort. It also should be pointed out that there was no significant difference in the *FOXA1* expression in four patients at different stages in our validation cohort. These issues and the underlying molecular mechanism deserve further study in the future.

In conclusion, the integrated analysis of miRNome profiling of human normal tissues and CRC tissues was performed to explore the roles of miRNAs in CRC in this study. In total, nine miRNAs (hsa-miR-492, hsa-miR-200a, hsa-miR-338, hsa-miR-29c, hsa-miR-101, hsa-miR-148a, hsa-miR-92a, hsa-miR-424, and hsa-miR-210) were proven to be promising diagnostic and prognostic markers in the clinic. Upregulation of miR-200a-3p suppressed the proliferation, migration, invasion, and EMT of colon cancer cells by targeting the *FOXA1/YAP1* signal pathway. Targeting miR-200a-3p might serve as a novel therapeutic approach for the treatment of CRC patients.

## DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found at The Cancer Genome Atlas and Gene Expression Omnibus: GSE48267 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE48267>); GSE38389 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE38389>); GSE-28364 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE28364>); GSE49246 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE49246>); GSE115513 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE115513>); GSE29622 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE29622>); and TC GA-COAD (<https://portal.gdc.cancer.gov/>).

## ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Research Ethics Committee of Hubei University of Medicine (Shiyan, Hubei, China).

The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

ZD, MD, WF, and QT participated in the research design. YL, PL, TL, and MS performed the data analysis. TL, XG, and MS wrote or contributed to the writing of the manuscript. All authors read and approved the final version of the manuscript.

## FUNDING

This research was supported by the National Natural Science Foundation of China (81902498), Natural Science Foundation of Hubei Province of China (2019CFB177), Natural Science Foundation of Hubei Provincial Department of Education (Q20182105), Chen Xiao-ping Foundation for the development of Science and Technology of Hubei Province (CXPJH11800001-2018333), Faculty Development Foundation of Hubei University of Medicine (2014QDJZR01), and Innovation and Entrepreneurship Training Program (201810929005, 201810929009, 201810929068, 201813249010, S201910929009, and S201910929045).

## ACKNOWLEDGMENTS

The authors would like to thank Medjaden Bioscience Limited. The authors would also like to thank the Cancer Genome

Atlas and the Gene Expression Omnibus for providing the data for this study.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2020.00192/full#supplementary-material>

**FIGURE S1** | Cross-validation for tuning parameter selection in the LASSO logistical model (A) and SVM-RFE model (B).

**FIGURE S2** | Establishment and validation of the composite clinicopathologic-miRNA nomogram for OS prediction. Nomograms for OS prediction in (A) testing cohort GSE29622 and (B) validation cohort TCGA-COAD. Calibration curve for the OS nomogram model in (C) testing cohort GSE29622, and (D) validation cohort TCGA-COAD. The dashed line represents the ideal nomogram, and the solid line represents the observed nomogram.

**FIGURE S3** | Scatter plots indicating the association between miR-200a and FOXA1 mRNA expression levels in (A) TCGA-COAD and (B) TCGA-Rectal cancer.

**TABLE S1** | Characteristics of the public microarray datasets used in this study.

**TABLE S2** | MiRNA-level results in integrated analysis of three datasets.

**TABLE S3** | Differentially expressed miRNAs between CRC and ANT selected using the LASSO and SVM-RFE algorithms.

**TABLE S4** | Clinical features of the CRC patients in the testing set (GSE29622), validation set (TCGA-COAD), and independent validation set.

**TABLE S5** | Sixty-six consensus target genes summarized by four different target prediction algorithms.

**DATA SHEET S1** | Supplementary experimental procedures.

## REFERENCES

- Agesen, T. H., Sveen, A., Merok, M. A., Lind, G. E., Nesbakken, A., Skotheim, R. I., et al. (2012). ColoGuideEx: a robust gene classifier specific for stage II colorectal cancer prognosis. *Gut* 61, 1560–1567.
- Alcantara, K. M. M., and Garcia, R. L. (2019). MicroRNA92a promotes cell proliferation, migration and survival by directly targeting the tumor suppressor gene NF2 in colorectal and lung cancer cells. *Oncol. Rep.* 41, 2103–2116. doi: 10.3892/or.2019.7020
- Balaguer, F., Moreira, L., Lozano, J. J., Link, A., Ramirez, G., Shen, Y., et al. (2011). Colorectal cancers with microsatellite instability display unique miRNA profiles. *Clin. Cancer Res.* 17, 6239–6249. doi: 10.1158/1078-0432.CCR-11-1424
- Chen, D. T., Hernandez, J. M., Shibata, D., McCarthy, S. M., Humphries, L. A., Clark, W., et al. (2012). Complementary strand microRNAs mediate acquisition of metastatic potential in colonic adenocarcinoma. *J. Gastrointest. Surg.* 16, 905–912. doi: 10.1007/s11605-011-1815-0 discussion 912–913
- Chen, E., Li, Q., Wang, H., Yang, F., Min, L., and Yang, J. (2018). MiR-92a promotes tumorigenesis of colorectal cancer, a transcriptomic and functional based study. *Biomed. Pharmacother.* 106, 1370–1377. doi: 10.1016/j.biopha.2018.07.098
- Chen, X., Liu, K., Yang, P., Kuang, W., Huang, H., Tu, E., et al. (2019). microRNA-200a functions as a tumor suppressor by targeting FOXA1 in glioma. *Exp. Ther. Med.* 17, 221–229. doi: 10.3892/etm.2018.6895
- Chen, X., Shi, K., Wang, Y., Song, M., Zhou, W., Tu, H., et al. (2015). Clinical value of integrated-signature miRNAs in colorectal cancer: miRNA expression profiling analysis and experimental validation. *Oncotarget* 6, 37544–37556. doi: 10.18632/oncotarget.6065
- Danielsen, H. E., Hveem, T. S., Domingo, E., Pradhan, M., Kleppe, A., Syvertsen, R. A., et al. (2018). Prognostic markers for colorectal cancer: estimating ploidy and stroma. *Ann. Oncol.* 29, 616–623. doi: 10.1093/annonc/mdx794
- Dhawan, A., Scott, J. G., Harris, A. L., and Buffa, F. M. (2018). Pan-cancer characterisation of microRNA across cancer hallmarks reveals microRNA-mediated downregulation of tumour suppressors. *Nat. Commun.* 9:5228. doi: 10.1038/s41467-018-07657-1
- Fitamant, J., Kottakis, F., Benhamouche, S., Tian, H. S., Chuvin, N., Parachoniak, C. A., et al. (2015). YAP inhibition restores hepatocyte differentiation in advanced HCC. Leading to tumor regression. *Cell Rep.* 10, 1692–1707. doi: 10.1016/j.celrep.2015.02.027
- Fu, F., Jiang, W., Zhou, L., and Chen, Z. (2018). Circulating exosomal miR-17-5p and miR-92a-3p predict pathologic stage and grade of colorectal cancer. *Transl. Oncol.* 11, 221–232. doi: 10.1016/j.tranon.2017.12.012
- Gaedcke, J., Grade, M., Camps, J., Sokilde, R., Kaczowski, B., Schetter, A. J., et al. (2012). The rectal cancer microRNAome-microRNA expression in rectal cancer and matched normal mucosa. *Clin. Cancer Res.* 18, 4919–4930.
- Gangadhar, T., and Schilsky, R. L. (2010). Molecular markers to individualize adjuvant therapy for colon cancer. *Nat. Rev. Clin. Oncol.* 7, 318–325. doi: 10.1038/nrclinonc.2010.62
- Huang, B., Yang, H., Cheng, X., Wang, D., Fu, S., Shen, W., et al. (2017). tRF/miR-1280 suppresses stem cell-like cells and metastasis in colorectal cancer. *Cancer Res.* 77, 3194–3206. doi: 10.1158/0008-5472.CAN-16-3146
- Hur, K., Toiyama, Y., Schetter, A. J., Okugawa, Y., Harris, C. C., Boland, C. R., et al. (2015). Identification of a metastasis-specific MicroRNA signature in human colorectal cancer. *J. Natl. Cancer Inst.* 107:dju492. doi: 10.1093/jnci/dju492

- Hur, K., Toiyama, Y., Takahashi, M., Balaguer, F., Nagasaka, T., Koike, J., et al. (2013). MicroRNA-200c modulates epithelial-to-mesenchymal transition (EMT) in human colorectal cancer metastasis. *Gut* 62, 1315–1326. doi: 10.1136/gutjnl-2011-301846
- Igder, S., Mohammadiasl, J., and Mokarram, P. (2019). Altered miR-21, miRNA-148a expression in relation to KRAS mutation status as indicator of adenoma-carcinoma transitional pattern in colorectal adenoma and carcinoma lesions. *Biochem. Genet.* 57, 767–780. doi: 10.1007/s10528-019-09918-0
- Ji, D., Qiao, M., Yao, Y., Li, M., Chen, H., Dong, Q., et al. (2018). Serum-based microRNA signature predicts relapse and therapeutic outcome of adjuvant chemotherapy in colorectal cancer patients. *EBioMedicine* 35, 189–197. doi: 10.1016/j.ebiom.2018.08.042
- Kandimalla, R., Gao, F., Matsuyama, T., Ishikawa, T., Uetake, H., Takahashi, N., et al. (2018). Genome-wide discovery and identification of a novel miRNA signature for recurrence prediction in stage II and III colorectal cancer. *Clin. Cancer Res.* 24, 3867–3877. doi: 10.1158/1078-0432.CCR-17-3236
- Kennedy, R. D., Bylesjo, M., Kerr, P., Davison, T., Black, J. M., Kay, E. W., et al. (2011). Development and independent validation of a prognostic assay for stage II colon cancer using formalin-fixed paraffin-embedded tissue. *J. Clin. Oncol.* 29, 4620–4626. doi: 10.1200/JCO.2011.35.4498
- Kim, J., Lim, N. J., Jang, S. G., Kim, H. K., and Lee, G. K. (2014). miR-592 and miR-552 can distinguish between primary lung adenocarcinoma and colorectal cancer metastases in the lung. *Anticancer Res.* 34, 2297–2302.
- Kuo, T. Y., Hsi, E., Yang, I. P., Tsai, P. C., Wang, J. Y., and Juo, S. H. (2012). Computational analysis of mRNA expression profiles identifies microRNA-29a/c as predictor of colorectal cancer early recurrence. *PLoS One* 7:e31587. doi: 10.1371/journal.pone.0031587
- Lin, X., Wang, S., Sun, M., Zhang, C., Wei, C., Yang, C., et al. (2019). miR-195-5p/NOTCH2-mediated EMT modulates IL-4 secretion in colorectal cancer to affect M2-like TAM polarization. *J. Hematol. Oncol.* 12:20. doi: 10.1186/s13045-019-0708-7
- Linnekamp, J. F., Wang, X., Medema, J. P., and Vermeulen, L. (2015). Colorectal cancer heterogeneity and targeted therapy: a case for molecular disease subtypes. *Cancer Res.* 75, 245–249. doi: 10.1158/0008-5472.CAN-14-2240
- Ma, W., Jiang, J., Li, M., Wang, H., Zhang, H., He, X., et al. (2016). The clinical significance of forkhead box protein?A1 and its role in colorectal cancer. *Mol. Med. Rep.* 14, 2625–2631. doi: 10.3892/mmr.2016.5583
- Ma, Y., Zhang, P., Wang, F., Zhang, H., Yang, J., Peng, J., et al. (2012). miR-150 as a potential biomarker associated with prognosis and therapeutic outcome in colorectal cancer. *Gut* 61, 1447–1453. doi: 10.1136/gutjnl-2011-301122
- Manceau, G., Imbeaud, S., Thiebaut, R., Liebaert, F., Fontaine, K., Rousseau, F., et al. (2014). Hsa-miR-31-3p expression is linked to progression-free survival in patients with KRAS wild-type metastatic colorectal cancer treated with anti-EGFR therapy. *Clin. Cancer Res.* 20, 3338–3347. doi: 10.1158/1078-0432.ccr-13-2750
- Marisa, L., de Reynies, A., Duval, A., Selves, J., Gaub, M. P., Vescovo, L., et al. (2013). Gene expression classification of colon cancer into molecular subtypes: characterization, validation, and prognostic value. *PLoS Med.* 10:e1001453. doi: 10.1371/journal.pmed.1001453
- Ozawa, T., Kandimalla, R., Gao, F., Nozawa, H., Hata, K., Nagata, H., et al. (2018). A MicroRNA signature associated with metastasis of T1 colorectal cancers to lymph nodes. *Gastroenterology* 154, 844–848e847. doi: 10.1053/j.gastro.2017.11.275
- Pettit, C., Walston, S., Wald, P., Webb, A., and Williams, T. M. (2017). Molecular profiling of locally-advanced rectal adenocarcinoma using microRNA expression (Review). *Int. J. Oncol.* 51, 393–404. doi: 10.3892/ijo.2017.4045
- Pichler, M., Ress, A. L., Winter, E., Stiegelbauer, V., Karbiener, M., Schwarzenbacher, D., et al. (2014). MiR-200a regulates epithelial to mesenchymal transition-related gene expression and determines prognosis in colorectal cancer patients. *Br. J. Cancer* 110, 1614–1621. doi: 10.1038/bjc.2014.51
- Qiu, J., Peng, B., Tang, Y., Qian, Y., Guo, P., Li, M., et al. (2017). CpG methylation signature predicts recurrence in early-stage hepatocellular carcinoma: results from a multicenter study. *J. Clin. Oncol.* 35, 734–742. doi: 10.1200/JCO.2016.68.2153
- Ramasamy, A., Mondry, A., Holmes, C. C., and Altman, D. G. (2008). Key issues in conducting a meta-analysis of gene expression microarray datasets. *PLoS Med.* 5:e184. doi: 10.1371/journal.pmed.0050184
- Ribic, C. M., Sargent, D. J., Moore, M. J., Thibodeau, S. N., French, A. J., Goldberg, R. M., et al. (2003). Tumor microsatellite-instability status as a predictor of benefit from fluorouracil-based adjuvant chemotherapy for colon cancer. *N. Engl. J. Med.* 349, 247–257. doi: 10.1056/NEJMoa022289
- Schee, K., Boye, K., Abrahamsen, T. W., Fodstad, Y., and Flatmark, K. (2012). 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
- Siegel, R. L., Miller, K. D., Fedewa, S. A., Ahnen, D. J., Meester, R. G. S., Barzi, A., et al. (2017). Colorectal cancer statistics, 2017. *CA Cancer J. Clin.* 67, 177–193. doi: 10.3322/caac.21395
- Siegel, R. L., Miller, K. D., and Jemal, A. (2016). Cancer statistics, 2016. *CA Cancer J. Clin.* 66, 7–30. doi: 10.3322/caac.21332
- Smits, K. M., Paranjape, T., Nallur, S., Wouters, K. A., Weijenberg, M. P., Schouten, L. J., et al. (2011). A let-7 microRNA SNP in the KRAS 3'UTR is prognostic in early-stage colorectal cancer. *Clin. Cancer Res.* 17, 7723–7731. doi: 10.1158/1078-0432.CCR-11-0990
- Song, B., Park, S. H., Zhao, J. C., Fong, K. W., Li, S., Lee, Y., et al. (2019). Targeting FOXA1-mediated repression of TGF-beta signaling suppresses castration-resistant prostate cancer progression. *J. Clin. Invest.* 129, 569–582. doi: 10.1172/jci122367
- Sun, K., Su, G., Deng, H., Dong, J., Lei, S., and Li, G. (2014). Relationship between miRNA-338-3p expression and progression and prognosis of human colorectal carcinoma. *Chin. Med. J.* 127, 1884–1890.
- Sun, M., Song, H., Wang, S., Zhang, C., Zheng, L., Chen, F., et al. (2017a). Integrated analysis identifies microRNA-195 as a suppressor of Hippo-YAP pathway in colorectal cancer. *J. Hematol. Oncol.* 10:79. doi: 10.1186/s13045-017-0445-8
- Sun, M., Sun, T., He, Z., and Xiong, B. (2017b). Identification of two novel biomarkers of rectal carcinoma progression and prognosis via co-expression network analysis. *Oncotarget* 8, 69594–69609. doi: 10.18632/oncotarget.18646
- Toiyama, Y., Okugawa, Y., Tanaka, K., Araki, T., Uchida, K., Hishida, A., et al. (2017). A panel of methylated microRNA biomarkers for identifying high-risk patients with ulcerative colitis-associated colorectal cancer. *Gastroenterology* 153, 1634–1646.e8. doi: 10.1053/j.gastro.2017.08.037
- Tsai, H. L., Yang, I. P., Huang, C. W., Ma, C. J., Kuo, C. H., Lu, C. Y., et al. (2013). Clinical significance of microRNA-148a in patients with early relapse of stage II stage and III colorectal cancer after curative resection. *Transl. Res.* 162, 258–268. doi: 10.1016/j.trsl.2013.07.009
- Tseng, G. C., Ghosh, D., and Feingold, E. (2012). Comprehensive literature review and statistical considerations for microarray meta-analysis. *Nucleic Acids Res.* 40, 3785–3799. doi: 10.1093/nar/gkr1265
- Vychytilova-Faltejskova, P., Radova, L., Sachlova, M., Kosarova, Z., Slaba, K., Fabian, P., et al. (2016). Serum-based microRNA signatures in early diagnosis and prognosis prediction of colon cancer. *Carcinogenesis* 37, 941–950. doi: 10.1093/carcin/bgw078
- Wang, S., Liu, F., Wang, Y., Fan, W., Zhao, H., Liu, L., et al. (2019). Integrated analysis of 34 microarray datasets reveals CBX3 as a diagnostic and prognostic biomarker in glioblastoma. *J. Transl. Med.* 17:179. doi: 10.1186/s12967-019-1930-3
- Wang, X., Kang, D. D., Shen, K., Song, C., Lu, S., Chang, L. C., et al. (2012). An R package suite for microarray meta-analysis in quality control, differentially expressed gene analysis and pathway enrichment detection. *Bioinformatics* 28, 2534–2536. doi: 10.1093/bioinformatics/bts485
- Wu, C. W., Ng, S. C., Dong, Y., Tian, L., Ng, S. S., Leung, W. W., et al. (2014). Identification of microRNA-135b in stool as a potential noninvasive biomarker for colorectal cancer and adenoma. *Clin. Cancer Res.* 20, 2994–3002. doi: 10.1158/1078-0432.CCR-13-1750
- Xue, Q., Sun, K., Deng, H. J., Lei, S. T., Dong, J. Q., and Li, G. X. (2014). MicroRNA-338-3p inhibits colorectal carcinoma cell invasion and migration by targeting smoothed. *Jpn. J. Clin. Oncol.* 44, 13–21. doi: 10.1093/jjco/hyt181

- Yang, Q., Yu, W., and Han, X. (2019). Overexpression of microRNA101 causes antitumor effects by targeting CREB1 in colon cancer. *Mol. Med. Rep.* 19, 3159–3167. doi: 10.3892/mmr.2019.9952
- Yong, F. L., Law, C. W., and Wang, C. W. (2013). Potentiality of a triple microRNA classifier: miR-193a-3p, miR-23a and miR-338-5p for early detection of colorectal cancer. *BMC Cancer* 13:280. doi: 10.1186/1471-2407-13-280
- Zhang, J. X., Song, W., Chen, Z. H., Wei, J. H., Liao, Y. J., Lei, J., et al. (2013). Prognostic and predictive value of a microRNA signature in stage II colon cancer: a microRNA expression analysis. *Lancet Oncol.* 14, 1295–1306. doi: 10.1016/s1470-2045(13)70491-1

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Di, Di, Fu, Tang, Liu, Lei, Gu, Liu and Sun. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.