Carcinogenesis, 2016, Vol. 37, No. 3, 245–261

doi:10.1093/carcin/bgv249
Advance Access publication January 6, 2016
Original Manuscript

ORIGINAL MANUSCRIPT

MicroRNA profiles in colorectal carcinomas, adenomas and normal colonic mucosa: variations in miRNA expression and disease progression

Martha L.Slattery*, Jennifer S.Herrick, Daniel F.Pellatt, John R.Stevens[1], Lila E.Mullany, Erica Wolff, Michael D.Hoffman, Wade S.Samowitz[2] and Roger K.Wolff

Department of Medicine, University of Utah, 383 Colorow, Salt Lake City, UT 84108, USA, [1]Department of Mathematics and Statistics, Utah State University, 3900 Old Main Hill, Logan, UT 84322-3900, USA and [2]Department of Pathology, University of Utah School, Salt Lake City, UT 84108, USA

*To whom correspondence should be addressed. Tel: 801-585-6955; Email: Marty.Slattery@hsc.utah.edu

Abstract

MiRNAs are small, non-protein-coding RNA molecules that regulate gene expression either by post-transcriptionally suppressing mRNA translation or by mRNA degradation. We examine differentially expressed miRNAs in colorectal carcinomas, adenomas and normal colonic mucosa. Data come from population-based studies of colorectal cancer conducted in Utah and the Kaiser Permanente Medical Care Program. A total of 1893 carcinoma/normal-paired samples and 290 adenoma tissue samples were run on the Agilent Human miRNA Microarray V19.0 which contained 2006 miRNAs. We tested for significant differences in miRNA expression between paired carcinoma/adenoma/normal colonic tissue samples. Fewer than 600 miRNAs were expressed in >80% of people for colonic tissue; of these 86.5% were statistically differentially expressed between carcinoma and normal colonic mucosa using a false discovery rate of 0.05. Roughly half of these differentially expressed miRNAs showed a progression in levels of expression from normal to adenoma to carcinoma tissue. Other miRNAs appeared to be altered at the normal to adenoma stage, while others were only altered at the adenoma to carcinoma stage or only at the normal to carcinoma stage. Evaluation of the Agilent platform showed a high degree of repeatability ($r = 0.98$) and reasonable agreement with the NanoString platform. Our data suggest that miRNAs are highly dysregulated in colorectal tissue among individuals with colorectal cancer; the pattern of disruption varies by miRNA as tissue progresses from normal to adenoma to carcinoma.

Introduction

'Every cellular process is likely to be regulated by microRNAs, and an aberrant microRNA expression signature is a hallmark of several diseases, including cancer' (1). MiRNAs are small, non-protein-coding RNA molecules that regulate gene expression either by post-transcriptionally suppressing mRNA translation or by causing mRNA degradation (2–7). In 2003, miRNAs were first reported as being associated with colorectal cancer (8). At that time, Michael *et al.* (8) described carcinoma suppressor-like activity for miR-143 and miR-145 in colon cancers and hypothesized that these miRNAs were targeting *ERK5* and *IRS1*. Since then, numerous studies have shown that miRNA are extensively involved in cancer pathogenesis of solid carcinomas, including colorectal cancer (9,10). Although studies generally have focused on describing differential expression of miRNA between carcinoma and non-carcinoma tissue, some of these miRNAs are being studied in detail to gain insight into specific mechanisms and targets. We know that miRNAs play a critical role in regulation of proliferation, differentiation, apoptosis and stress response and are involved in the majority of physiological processes (1,11). Thus, it is important to obtain a better understanding of the role miRNAs play in the carcinogenic process.

Received: October 9, 2015; **Revised:** November 30, 2015; **Accepted:** December 11, 2015

Abbreviations

FDR	false discovery rate
KPMCP	Kaiser Permanente Medical Care Program

The goal of this study is to evaluate miRNAs as they relate to colon and rectal cancer. This includes discovery of new miRNAs that may contribute to cancer and prognosis as well as to replicate previous findings. We evaluate cancer progression by including miRNA expression profiles of almost 2000 cases that includes carcinoma and non-carcinoma/non-adenoma paired colorectal tissue as well as miRNA expression profiles for adenomas associated with 290 of the cases. In this study, we focus on miRNAs that are differentially expressed between carcinoma and non-carcinoma/non-adenoma colorectal tissue, adenomas and non-carcinoma/non-adenoma colorectal tissue, and carcinoma and adenoma colorectal tissue. Given our large sample size we are able to describe miRNA expression profiles in terms of level and frequency of expression. We address issues pertaining to platform selection, normalization techniques, nomenclature changes and statistical analysis that exist when doing a large discovery study that encompass thousands of miRNAs rather than select targeted miRNAs as is often reported.

Methods

Study participants

The study was approved by the Institutional Review Board of the University of Utah. Study participants came from two population-based case–control studies that included all incident colon and rectal cancers between 30 and 79 years of age who resided along the Wasatch Front in Utah or were members of the Kaiser Permanente Medical Care Program (KPMCP) in Northern California. Participants were white, Hispanic or black for the colon cancer study (diagnosed between October 1991 and September 1994); the rectal cancer study (diagnosed between June 1997 and May 2001) also included Asians and American Indians not living on reservations (12,13). Tumor tissue was obtained for 97% of all Utah cases diagnosed and for 85% of all KPMCP study participants (14).

Of the 2619 participants who were targeted for the study, we could not make miRNA on 637 because of too little tissue (Supplementary Table 1, available at *Carcinogenesis* Online). Both carcinoma and normal colonic mucosa (subsequently called normal) miRNA scans were obtained for 1657 individuals, carcinoma only miRNA scans for 297 participants, and normal only scans for 21 people. Both the carcinoma and normal tissue microarray failed for seven people. We targeted adenoma tissue for 388 individuals with carcinoma and obtained adenoma scans on 298 people. We could not make adenoma RNA for 84 adenomas and 6 microarrays failed. Since the study focuses on matched carcinoma–adenoma–normal samples, we excluded from analysis individuals whose microarray scan showed weak signal (i.e. 60 carcinomas, 80 normals and 2 adenomas) and those whose carcinoma could not be obtained (i.e. 59 individuals with normal tissue only when the carcinoma tissue microarray failed or did not pass QC and six individuals with adenoma only tissue when carcinoma tissue failed). After imputing normal miRNA (see Statistical methods for details) values for 354 individuals for those with only carcinoma tissue, we had a total of 1893 carcinoma/normal pairs and 290 individuals with carcinoma/adenoma/normal for analysis.

miRNA processing

RNA was extracted from formalin-fixed paraffin embedded tissue. We assessed slides and carcinoma blocks that were prepared over the duration of the study prior to the time of miRNA isolation to determine their suitability. The study pathologist (W.S.) reviewed slides to delineate carcinoma, normal, and adenoma tissue. Normal tissue adjacent to the carcinoma tissue was used. Cells were dissected from 1 to 4 sequential sections on aniline blue stained slides using an H&E slide for reference. Total RNA containing miRNA was extracted, isolated and purified using the RecoverAll Total Nucleic Acid isolation kit (Ambion), RNA yields were determined using a NanoDrop spectrophotometer.

The Agilent Human miRNA Microarray V19.0 was used given the number of miRNAs, its high level of reliability and the amount of RNA needed to run the platform. The microarray contains probes for 2006 unique human miRNAs. The miRNA array contains on average 30 replicates per probe sequence for a total of 60 000 unique features. The Agilent Human microarray was generated using known miRNA sequence information compiled in the Sanger miRBASE database v19.0. About 100 ng total RNA was labeled with Cy3 and hybridized to the Agilent Microarray and were scanned on an Agilent SureScan microarray scanner model G2600D. Data were extracted from the scanned image using Agilent Feature Extract software v.11.5.1.1. Data were required to pass stringent QC parameters established by Agilent that included tests for excessive background fluorescence, excessive variation among probe sequence replicates on the array, and measures of the total gene signal on the array to assess low signal. If samples failed to meet quality standards for any of these parameters, the sample was re-labeled, hybridized to arrays and scanned. If a sample failed QC assessment a second time the sample was deemed to be of poor quality and the individual was excluded from downstream analysis. To test for reliability of the Agilent Microarray over time, we repeated 13 samples (8 carcinoma and 5 matched normal), taking samples that had scans run over the course of the study.

We ran the NanoString Platform on 30 samples that had both carcinoma and normal Agilent Microarray data using RNA from the same prep that was used on the Agilent Microarray. Of these 30 samples analyzed with NanoString, we repeated five-matched carcinoma/normal paired samples to determine reliability of the platform.

miRNA nomenclature

We refer to miRNAs using standard nomenclature used in the miRBase database (15). Briefly, the first three letters signifies the organism, followed by a unique number. The number is followed by a dash and number (i.e. –1) if more than one locus codes for the miRNA. A lettered suffix denotes closely related miRNAs. If two miRNAs are coded by the same precursor product then the minor product is assigned the suffix (*). If predominant/minor product status is not known then the suffix –5p and –3p are used to denote 5′ and 3′ arm, respectively. Nomenclature changes in the literature and across platforms exist, creating difficulties in comparing results. For instance, let-7 may be reported in the literature as being associated with carcinoma stage, however Let-7 has since been further delineated to several closely related mature sequences and genomic loci, e.g. let-7a-3p, let-7a-5p and let-7b-3p.

Statistical methods

Imputation

Imputation of normal was done for those individuals where carcinoma tissue was available but we could not obtain normal tissue (103 proximal colon, 93 distal colon and 158 rectal). We used the methods as described by Suyundikov *et al*. for imputation (16). Imputation was done separately by carcinoma subsite (i.e. proximal or distal colon or rectal). The K-nearest neighbor method incorporates other information including age, stage, sex, study (i.e. colon versus rectal case–control study) and center (Utah or KPMCP). This method is highly predictive of actual measured values and relies on nearest 10 individuals who best match demographic and miRNA data of each missing-normal individual.

Normalization of miRNA expression

To minimize differences that could be attributed to the array, amount of RNA, location on array or other factors that could erroneously influence expression, total gene signal was normalized by multiplying each sample by a scaling factor stratified by carcinoma site for the Agilent Microarray Data. Those miRNAs with a value set to 0.1 were considered as not-expressing beyond

Table 1. Description of study population and miRNA expression

	Overall		Colon		Rectal	
	Subject *N*	%	Subject *N*	%	Subject *N*	%
Sex						
Male	1028	54.3	608	52.8	420	56.5
Female	866	45.7	543	47.2	323	43.5
Center						
Kaiser	1144	60.4	740	64.3	404	54.4
Utah	750	39.6	411	35.7	339	45.6
Site						
Proximal colon	569	49.5	569	49.4	0	0.0
Distal colon	580	50.5	580	50.4	0	0.0
Study						
Stage I	559	30.0	259	22.7	300	41.5
Stage II	489	26.3	350	30.7	139	19.2
Stage III	548	29.4	340	29.9	208	28.8
Stage IV	266	14.3	190	16.7	76	10.5
TP53						
Wt	953	52.4	597	54.4	356	49.4
Mut	864	47.6	500	45.6	364	50.6
KRAS						
Wt	1240	68.5	724	67.6	516	69.9
Mut	569	31.5	347	32.4	222	30.1
CIMP						
Low	1312	78.8	700	71.8	612	88.6
High	354	21.2	275	28.2	79	11.4
MSI						
Stable	1688	90.9	965	86.2	723	97.8
Unstable	170	9.1	154	13.8	16	2.2
	Mean	STD	Mean	STD	Mean	STD
Age	64.2	10.2	65.4	9.5	62.3	11.0
	Carcinoma		Adenoma		Normal	
miRNAs	miRNA *N*	%	miRNA *N*	%	miRNA *N*	%
Not expressed	728	36.29	934	46.56	737	36.74
Expressed	1278	63.71	1072	53.44	1269	63.26
Percent expressing						
0>–10	441	34.51	272	25.37	394	31.05
>10–20	57	4.46	47	4.38	63	4.96
>20–30	39	3.05	21	1.96	37	2.92
>30–40	39	3.05	28	2.61	36	2.84
>40–50	35	2.74	21	1.96	45	3.55
>50–60	21	1.64	29	2.71	35	2.76
>60–70	25	1.96	20	1.87	25	1.97
>70–80	33	2.58	38	3.54	39	3.07
>80–90	48	3.76	50	4.66	48	3.78
>90–100	540	42.25	546	50.93	547	43.1

background and were set to 0 expression. Within each carcinoma site, the scaling factor (17) (http://genespring-support.com/files/gs_12_6/GeneSpring-manual.pdf) was the median of the 75th percentiles of all the samples divided by the individual 75th percentile of each sample.

For the NanoString data, we calculated a background level of expression for each sample using the mean level of the negative controls plus two standard deviations of the mean. MiRNAs expressing less than two standard deviations from the mean were set to 0 expression. Those miRNAs that were considered non-zero expression, were normalized using a scaling factor based on the top 100 expressing miRNAs across all samples. For each sample, the average of the geometric means of the top 100 expressing miRNAs across all samples was divided by the geometric mean of each sample (http://www.NanoString.com/media/pdf/MAN_nCounter_Gene_Expression_Data_Analysis_Guidelines.pdf).

MiRNA statistical analysis

All analysis was based on individual's paired data. Given the size of the study, we are able to describe miRNAs that are commonly expressed as well as those that are less frequently expressed. Among miRNAs more commonly expressed (defined as expressed in over 20% of individuals and included 814 miRNAs), the data were randomly split into two groups. Within each group, differential expression between tissue types was

analyzed using the significance analysis of microarrays technique implemented in the R package siggenes (18); *P* values were based upon 1000 permutations. As a form of cross validation, this analysis was run in both groups and the larger *P* value was kept to take forward for adjustment of multiple comparisons to determine a false discovery rate (FDR) level of significance. For the group of miRNAs expressed in less than 20% of the population (580 miRNAs), we used a chi-squared test to determine if tissue type was independent of expression versus non-expression. We determined if miRNAs are differentially expressed, using log base 2 transformed miRNA expression levels, between carcinoma and normal mucosa, between adenoma tissue and normal mucosa and between carcinoma and adenoma tissue.

Multiple comparison adjustment took into account *P* values from those commonly expressed miRNAs as well as those from miRNAs less frequently expressed in order to account for all comparisons made. *P* values from the chi-squared analysis on miRNA expressed less frequently percentage of individuals were combined with the permuted *P* values from sigGenes and those significant when applying a FDR of 0.05 according to the Benjamini and Hochberg comparisons were identified (19). All analyses except significance analysis of microarrays were performed using SAS 9.4 (SAS Institute, Cary, NC). For those miRNAs that were significantly differentially expressed, we report the mean level of expression and the fold change (on non-log-transformed data) between the carcinoma/adenoma and normal tissue calculated as the ratio of the mean level of expression in the carcinoma (or adenoma) to the mean level of expression in the normal tissue.

Comparison of Agilent and NanoString platforms

We compared replicate samples on Agilent and NanoString to determine the reliability of each platform as well as to how they related to each other. To compare the two platforms, we analyzed each miRNA that was expressed in over 80% of the population on the Agilent MicroArray Platform to those that also are included on the NanoString platform (N = 150). For each such miRNA, we calculated the Pearson correlation coefficient between the expression measurements from the two platforms; to summarize these results a frequency distribution was generated. Additionally we evaluated platform agreement, using all of the 664 miRNAs that had the same nomenclature for assessment between the two platforms, and all 2006 miRNAs for Agilent and 798 miRNAs for NanoString platform for agreement within the same platform. Concordant agreement was defined as both platforms showing no expression, both platforms showed upregulation (based on carcinoma versus normal differential expression), or miRNAs exhibited downregulation on both platforms (20). Discordance was defined as miRNAs being up- or downregulated on one platform but having the opposite direction of association on the other. We also considered those expressed on one platform but not the other.

Results

The study population was 54.3% male and roughly equal proportion of colon proximal and distal tumors; the average age was 64.2 years (Table 1). Of the 2006 miRNAs on the Agilent Microarray platform, 737 (36.74%) were not expressed in normal colonic mucosa. Adenomas had the most miRNAs that were not expressed (934 or 46.56% of all miRNAs). Close to 20% of the miRNAs that were expressed in normal colonic mucosa had expression in less than 10% of the samples; this figure was slightly lower for adenomas (13.56%) and higher for carcinomas (21.98%). Roughly 540 (27%) miRNAs were commonly expressed in over 90% of the population based on expression in normal tissue. Approximately one to three percent of miRNAs were expressed in each 10% category from >10 to 20% to >80 to 90% of the cases which translates to roughly 16% of miRNAs (Table 1).

Comparison of carcinoma tissue to normal colonic mucosa

Most of the differentially expressed miRNAs were commonly expressed miRNAs (>80% of the normal tissue samples). In this group, 517 of the 598 miRNAs were differentially expressed; those differentially expressed with a fold change of ≤0.68 or ≥1.5 (8.7% of those dysregulated) are shown in Table 2 (Supplementary Table 2, available at *Carcinogenesis* Online shows all carcinoma/normal differentially expressed miRNAs). Roughly 2/3 of miRNAs were downregulated. Fifteen of these miRNAs were dysregulated in rectal carcinomas only and four were dysregulated in colon carcinomas only. Similar results were observed for those expressing in 20–80% of the population (Supplementary Table 2, available at *Carcinogenesis* Online has all dysregulated genes in this category), where over 70% of miRNAs were dysregulated. However of 154 that were dysregulated, 40% had a fold change of ≤0.68 or ≥1.5 (Table 2). Expression levels of miRNAs expressed in <20% of samples are shown in Table 3; 54% of miRNAs among those infrequently expressed were differentially expressed between carcinoma and normal tissue. Unlike the more commonly expressed miRNAs, 13 of these dysregulated miRNAs expressed in <20% of the population were seen only for colon carcinomas and six were dysregulated for rectal carcinomas only. In samples where expression was less common, the level of expression was generally low. However, miRNAs were still more likely to be down-regulated in carcinoma tissue relative to normal tissue. Of the 238 differentially expressed miRNAs, 25 were not expressed in either carcinoma or non-carcinoma tissue, and 132 (55.3%) had a fold change of ≤0.68 or ≥1.5 (Table 3 shows those with a fold change of <0.5 of >2.0). It is interesting to note that among those miRNAs very infrequently expressed in the population, the fold change is very large in some instances due to minimal expression in carcinoma or normal tissue making any fold change in expression very large, although absolute differences in expression are generally much less than among those where the miRNA is expressed more frequently. Differences in frequency of expression were not associated with any one carcinoma molecular phenotype (i.e. *TP53*, *Kras*, CIMP, MSI), rather each carcinoma phenotype had roughly the same proportion of the population non-expressing that miRNA (data not shown).

Comparison of adenoma to carcinoma tissue and to non-carcinoma/non-adenoma tissue

Comparison of miRNA expression across carcinoma, adenoma and normal tissue revealed several patterns (Table 4 for those miRNAS expressed in 80% or more of the population and Table 5 for those miRNAs expressed in 20–80% of the population where the fold change is either ≥1.5 or ≤0.68 and the FDR is set at 0.05; Supplementary Table 4, available at *Carcinogenesis* Online shows all dysregulated miRNAs when the FDR is set at 0.05). Among those miRNAs expressed in over 80% of that population that were differentially expressed between carcinoma and normal, roughly half also were differentially expressed between carcinoma and adenoma as well as between adenoma and normal tissue. This pattern of progression from the normal tissue to adenoma to carcinoma was the most common pattern observed, especially

Table 2. All miRNAs expressing in over 20% of samples that are differentially expressed between carcinoma tissue and normal mucosa where fold change is <0.68 or >1.5

	Carcinoma		Normal		C/N	
	% Expressing	Mean expression	% Expressing	Mean expression	SigGenes *P* value	Fold change
Downregulated	Expressing in >80% of samples					
hsa-miR-145-5p*	99.8	128.07	100.0	221.17	<0.0001	0.58
hsa-miR-150-5p*	87.6	12.49	98.6	36.51	<0.0001	0.34
hsa-miR-195-5p*	46.2	3.01	84.2	10.32	<0.0001	0.29
hsa-miR-215	95.7	40.63	99.0	61.93	<0.0001	0.66
hsa-miR-375	82.6	17.42	99.4	49.33	<0.0001	0.35
hsa-miR-4323	97.9	8.26	99.5	12.11	<0.0001	0.68
hsa-miR-451a*	72.6	15.56	88.8	31.80	<0.0001	0.49
hsa-miR-4539	100.0	50.14	100.0	77.06	<0.0001	0.65
hsa-miR-4749-3p	95.9	8.83	99.0	13.42	<0.0001	0.66
hsa-miR-6073	84.8	4.61	93.3	6.76	<0.0001	0.68
hsa-miR-650	77.7	4.11	99.4	16.94	<0.0001	0.24
Upregulated						
hsa-let-7i-5p*	98.6	52.61	98.6	33.33	<0.0001	1.58
hsa-miR-10a-5p*	95.6	36.77	97.7	21.70	<0.0001	1.69
hsa-miR-148a-3p*	79.7	15.86	81.6	9.80	<0.0001	1.62
hsa-miR-151a-5p	91.8	13.54	90.4	8.80	<0.0001	1.54
hsa-miR-151b	83.9	6.57	80.3	4.33	<0.0001	1.52
hsa-miR-17-5p*	98.5	51.18	95.9	13.97	<0.0001	3.66
hsa-miR-193b-3p	86.8	8.71	83.1	4.93	<0.0001	1.76
hsa-miR-199a-3p	95.3	37.52	94.2	18.51	<0.0001	2.03
hsa-miR-199a-5p*	90.3	16.79	85.6	7.34	<0.0001	2.29
hsa-miR-19b-3p	93.1	24.07	83.8	7.79	<0.0001	3.09
hsa-miR-20a-5p*	97.8	58.90	94.6	14.83	<0.0001	3.97
hsa-miR-210	97.6	23.33	98.3	15.39	<0.0001	1.52
hsa-miR-21-3p*	96.5	19.74	87.0	8.51	<0.0001	2.32
hsa-miR-214-3p*	90.9	12.37	85.8	5.57	<0.0001	2.22
hsa-miR-21-5p*	99.5	386.81	99.5	130.75	<0.0001	2.96
hsa-miR-222-3p*	97.6	16.47	96.3	8.87	<0.0001	1.86
hsa-miR-23a-3p*	99.4	153.93	99.5	73.88	<0.0001	2.08
hsa-miR-24-3p	99.6	92.99	99.7	54.04	<0.0001	1.72
hsa-miR-25-3p*	95.9	25.37	89.8	10.07	<0.0001	2.52
hsa-miR-27a-3p*	98.1	48.02	96.2	18.55	<0.0001	2.59
hsa-miR-29a-3p*	98.7	92.41	99.1	40.15	<0.0001	2.30
hsa-miR-29b-3p*	94.4	19.98	89.7	7.56	<0.0001	2.64
hsa-miR-331-3p*	95.9	13.31	93.2	8.09	<0.0001	1.65
hsa-miR-34a-5p*	96.1	20.58	91.2	10.18	<0.0001	2.02
hsa-miR-361-5p	88.6	10.62	81.6	5.17	<0.0001	2.05
hsa-miR-3651	99.3	55.29	99.1	23.83	<0.0001	2.32
hsa-miR-424-3p*	99.9	36.48	99.9	23.97	<0.0001	1.52
hsa-miR-425-5p*	83.6	9.86	80.3	5.50	<0.0001	1.79
hsa-miR-4506	95.7	9.73	89.1	6.44	<0.0001	1.51
hsa-miR-501-3p	96.0	6.82	84.8	2.58	<0.0001	2.64

Table 2. *Continued*

	Carcinoma		Normal		C/N	
	% Expressing	Mean expression	% Expressing	Mean expression	SigGenes *P* value	Fold change
hsa-miR-663a*	100.0	397.96	100.0	261.63	<0.0001	1.52
hsa-miR-663b*	100.0	65.94	100.0	32.76	<0.0001	2.01
hsa-miR-92a-3p*	99.8	105.94	99.6	39.03	<0.0001	2.71
hsa-miR-93-5p*	97.8	35.04	94.6	12.51	<0.0001	2.80
Downregulated	Expressing in 20 of 80% of normal					
hsa-miR-1203	43.1	1.65	64.6	2.58	<0.0001	0.64
hsa-miR-1207-3p	14.1	0.79	34.2	1.42	<0.0001	0.56
hsa-miR-124-3p	17.1	1.02	36.3	1.60	<0.0001	0.64
hsa-miR-1258	27.1	2.11	48.2	3.38	<0.0001	0.62
hsa-miR-1271-5p	10.1	0.94	24.0	1.46	<0.0001	0.64
hsa-miR-133b*	17.0	1.63	66.1	5.98	<0.0001	0.27
hsa-miR-142-3p	30.8	2.54	57.5	4.64	<0.0001	0.55
hsa-miR-192-3p*	23.1	1.61	48.1	2.69	<0.0001	0.60
hsa-miR-204-3p*	11.4	0.89	31.8	1.62	<0.0001	0.55
hsa-miR-2117	25.8	1.50	54.0	3.46	<0.0001	0.43
hsa-miR-30a-5p	31.0	1.82	59.3	3.81	<0.0001	0.48
hsa-miR-3178	25.4	1.17	44.3	1.79	<0.0001	0.65
hsa-miR-3181	27.1	2.20	48.7	3.49	<0.0001	0.63
hsa-miR-330-3p*	50.9	3.26	74.0	5.55	<0.0001	0.59
hsa-miR-3607-5p	7.1	0.90	22.9	1.35	<0.0001	0.67
hsa-miR-378d	4.3	0.23	34.7	1.63	<0.0001	0.14
hsa-miR-378g	23.6	1.30	41.6	1.97	<0.0001	0.66
hsa-miR-3923	6.1	0.55	20.7	1.31	<0.0001	0.42
hsa-miR-4315	1.2	0.12	23.5	1.82	<0.0001	0.06
hsa-miR-4421	24.0	1.50	41.9	2.42	<0.0001	0.62
hsa-miR-4458	55.5	3.64	75.2	5.87	<0.0001	0.62
hsa-miR-4469	24.7	1.43	45.7	2.26	<0.0001	0.63
hsa-miR-4520b-3p	24.7	1.73	42.6	2.59	<0.0001	0.67
hsa-miR-497-5p	22.5	1.27	71.6	5.83	<0.0001	0.22
hsa-miR-513a-3p	9.3	0.58	24.2	1.08	<0.0001	0.54
hsa-miR-513c-3p	26.6	2.15	44.8	3.27	<0.0001	0.66
hsa-miR-5685	19.4	1.55	42.9	2.67	<0.0001	0.58
hsa-miR-6071	15.0	0.95	37.0	1.65	<0.0001	0.57
hsa-miR-6515-5p	27.5	1.17	79.2	5.01	<0.0001	0.23
hsa-miR-659-5p	23.7	2.02	42.0	2.99	<0.0001	0.68
Upregulated						
hsa-miR-106b-5p*	88.5	13.12	71.3	3.95	<0.0001	3.32
hsa-miR-1291	76.3	5.74	69.1	3.29	<0.0001	1.74
hsa-miR-130b-3p*	85.7	7.91	72.1	4.23	<0.0001	1.87
hsa-miR-146a-5p	70.2	8.17	73.5	5.32	0.0005	1.54
hsa-miR-146b-5p	45.0	3.48	43.1	1.78	<0.0001	1.95
hsa-miR-151a-3p*	51.6	4.42	33.8	1.10	<0.0001	4.01
hsa-miR-15a-5p	62.8	6.26	62.9	3.68	<0.0001	1.70

Table 2. *Continued*

	Carcinoma		Normal		C/N	
	% Expressing	Mean expression	% Expressing	Mean expression	SigGenes *P* value	Fold change
hsa-miR-196a-5p	63.7	6.19	66.9	3.32	<0.0001	1.87
hsa-miR-196b-5p	76.8	15.32	71.0	4.73	<0.0001	3.24
hsa-miR-199b-5p*	39.4	3.58	32.5	1.18	<0.0001	3.05
hsa-miR-203a*	69.6	11.68	46.8	2.67	<0.0001	4.37
hsa-miR-20b-5p*	88.6	14.42	61.0	2.53	<0.0001	5.70
hsa-miR-221-3p*	85.4	11.40	63.7	2.74	<0.0001	4.16
hsa-miR-324-5p*	39.7	4.23	34.0	1.79	<0.0001	2.36
hsa-miR-3622b-3p	34.6	2.50	20.9	0.84	<0.0001	2.98
hsa-miR-365a-3p*	65.4	7.75	51.2	3.45	<0.0001	2.24
hsa-miR-3677-3p	59.9	4.84	52.4	2.98	<0.0001	1.63
hsa-miR-3972	47.5	3.91	47.9	2.33	<0.0001	1.68
hsa-miR-3976	30.1	2.28	27.6	1.03	<0.0001	2.22
hsa-miR-4251	39.2	3.25	32.6	1.45	<0.0001	2.24
hsa-miR-429	65.9	11.33	65.4	6.01	<0.0001	1.89
hsa-miR-4296	36.9	1.86	37.2	1.23	<0.0001	1.51
hsa-miR-4317	32.3	2.07	24.4	1.00	<0.0001	2.07
hsa-miR-466	21.0	3.14	27.6	2.07	0.01	1.51
hsa-miR-4700-5p	23.9	1.26	26.1	0.84	0.0006	1.51
hsa-miR-483-3p*	43.2	9.00	48.9	4.27	0.02	2.11
hsa-miR-5008-3p	22.1	2.09	21.3	0.79	<0.0001	2.65
hsa-miR-532-3p	31.1	2.67	25.9	1.15	<0.0001	2.33
hsa-miR-583	87.4	7.17	71.9	3.42	<0.0001	2.09
hsa-miR-934	83.0	4.63	45.5	0.75	<0.0001	6.15
hsa-miR-99a-5p*	52.5	5.71	54.9	3.19	<0.0001	1.79
hsa-miR-99b-5p*	41.9	4.85	44.7	3.02	0.003[a]	1.61

[a]miRNA significant only in colon study.

Table 3. MiRNAs differentially expressed between carcinoma tissue and normal mucosa among those miRNAs expressed in <20% of the population

	Carcinoma		Normal			C/N
miRNA	% Expressing	Mean expression	% Expressing	Mean expression	SigGenes *P* value	Fold change
Downregulated						
hsa-miR-378a-5p	0.1	0.04	3.5	2.27	<0.0001	0.02
hsa-miR-342-5p*	0.1	0.12	0.5	3.24	0.01	0.04
hsa-miR-4764-5p	0.1	0.02	4.0	0.41	<0.0001	0.05
hsa-miR-29c-5p*	0.1	0.05	3.3	0.99	<0.0001	0.05
hsa-miR-142-5p	0.1	0.20	1.2	3.33	<0.0001	0.06
hsa-miR-4446-5p	0.1	0.28	2.6	4.18	<0.0001	0.07
hsa-miR-4658	0.2	0.09	3.0	1.21	<0.0001	0.07
hsa-miR-4729	0.1	0.17	1.1	2.06	<0.0001	0.08
hsa-miR-3116	0.2	0.26	1.1	2.24	0.0003	0.11
hsa-miR-2116-3p	0.5	0.11	10.2	0.76	<0.0001	0.14
hsa-miR-589-5p	0.1	0.21	0.6	1.49	0.003	0.14
hsa-miR-3192	0.4	1.37	1.4	8.91	0.002	0.15
hsa-miR-491-3p*	0.5	0.38	3.1	2.27	<0.0001	0.17
hsa-miR-3908	1.0	0.14	11.7	0.84	<0.0001	0.17
hsa-miR-5571-5p	1.8	0.31	14.1	1.72	<0.0001	0.18
hsa-miR-491-5p	0.3	0.08	2.9	0.45	<0.0001	0.19
hsa-miR-5591-5p	0.1	0.06	1.7	0.31	<0.0001	0.20
hsa-miR-4796-3p	0.8	0.13	6.7	0.61	<0.0001	0.22
hsa-miR-138-5p*	0.1	0.26	0.7	1.04	0.004	0.25
hsa-miR-4279	1.1	0.22	11.5	0.83	<0.0001	0.26
hsa-miR-133a*	2.0	0.47	12.4	1.64	<0.0001	0.29
hsa-miR-3065-3p	0.5	0.64	1.3	2.20	0.02	0.29
hsa-miR-126-5p*	0.4	0.24	2.2	0.80	<0.0001	0.30
hsa-miR-1	2.3	0.58	12.8	1.73	<0.0001	0.33
hsa-miR-4330	0.6	0.13	6.0	0.39	<0.0001	0.34
hsa-miR-5006-3p	0.1	0.33	0.7	0.97	0.004	0.34
hsa-miR-4258	0.3	0.18	2.2	0.54	<0.0001	0.34
hsa-miR-4252	0.9	0.17	7.0	0.50	<0.0001	0.34
hsa-miR-6719-3p	1.3	0.52	10.5	1.51	<0.0001	0.35
hsa-miR-4256	1.1	0.27	7.6	0.73	<0.0001	0.37
hsa-miR-3194-3p	1.1	0.24	6.0	0.64	<0.0001	0.38
hsa-miR-4754	1.1	0.52	3.6	1.35	<0.0001	0.39
hsa-miR-297	0.6	0.36	2.3	0.85	0.0009[a]	0.43
hsa-miR-20b-3p*	0.8	0.15	3.5	0.35	<0.0001	0.43
hsa-miR-4326	1.8	0.30	9.5	0.68	<0.0001	0.43
hsa-miR-3159	1.1	0.53	3.0	1.23	<0.0001	0.43
hsa-miR-4783-5p	0.6	0.23	4.5	0.50	<0.0001	0.45
hsa-miR-18b-3p*	0.2	0.31	3.3	0.67	<0.0001	0.46
hsa-miR-6503-3p	0.4	0.28	2.1	0.60	<0.0001	0.46
hsa-miR-3064-5p	2.6	0.43	11.7	0.91	<0.0001	0.47

Table 3. *Continued*

	Carcinoma		Normal			C/N
miRNA	% Expressing	Mean expression	% Expressing	Mean expression	SigGenes *P* value	Fold change
hsa-miR-301b	0.1	0.08	1.5	0.17	<0.0001	0.48
hsa-miR-145-3p*	1.7	0.55	9.3	1.15	<0.0001	0.48
hsa-miR-374a-3p	0.1	0.24	0.8	0.50	0.02[a]	0.48
Upregulated						
hsa-miR-1304-5p	0.7	0.59	1.8	0.26	0.003	2.27
hsa-miR-675-5p	8.4	0.69	12.8	0.29	<0.0001	2.40
hsa-miR-373-3p*	1.2	3.09	4.9	1.26	<0.0001	2.46
hsa-miR-204-5p*	4.5	3.60	8.2	1.44	<0.0001	2.50
hsa-miR-885-3p	2.3	0.29	4.3	0.11	0.03[b]	2.56
hsa-miR-5193	0.7	1.36	1.8	0.53	0.001	2.57
hsa-miR-652-3p	17.5	2.42	13.0	0.89	0.003[a]	2.73
hsa-miR-4426	1.6	0.28	4.0	0.10	0.005[b]	2.79
hsa-miR-127-3p*	13.7	2.60	10.4	0.92	0.02[a]	2.83
hsa-miR-744-5p	0.7	0.97	1.6	0.22	0.006	4.42
hsa-miR-185-5p	12.7	2.42	7.9	0.53	<0.0001	4.57
hsa-miR-196b-3p	0.6	3.65	0.1	0.76	0.006	4.81
hsa-miR-372*	0.1	5.03	1.5	1.03	<0.0001	4.88
hsa-miR-1248	0.9	1.65	0.2	0.32	0.003	5.09
hsa-miR-98-5p*	33.2	3.16	17.2	0.62	<0.0001	5.11
hsa-miR-532-5p*	40.0	5.05	19.8	0.91	<0.0001	5.54
hsa-miR-339-5p	8.8	0.82	6.3	0.15	0.004	5.54
hsa-miR-3180	7.2	1.54	3.2	0.27	<0.0001	5.78
hsa-miR-92b-5p*	1.1	0.58	0.2	0.10	0.0005	5.85
hsa-miR-29b-1-5p*	2.4	1.21	0.3	0.19	<0.0001	6.30
hsa-miR-362-5p*	33.4	3.18	15.6	0.48	<0.0001	6.58
hsa-miR-4422	1.2	7.47	0.2	1.09	0.0005	6.83
hsa-miR-374b-5p	29.8	5.98	11.7	0.84	<0.0001	7.15
hsa-miR-505-3p*	5.2	2.12	2.3	0.29	<0.0001	7.33
hsa-miR-424-5p*	35.4	3.65	14.3	0.48	<0.0001	7.65
hsa-miR-193a-3p	33.0	3.48	13.6	0.44	<0.0001	7.87
hsa-miR-3687	33.8	3.16	10.7	0.38	<0.0001	8.25
hsa-miR-6715b-3p	8.5	2.15	3.9	0.26	<0.0001	8.25
hsa-miR-4427	2.1	3.89	0.4	0.45	<0.0001	8.70
hsa-miR-128*	7.6	2.49	2.7	0.28	<0.0001	8.94
hsa-miR-34b-5p*	1.7	3.34	0.3	0.36	<0.0001	9.27
hsa-miR-660-5p	11.2	3.20	2.0	0.34	<0.0001	9.28
hsa-miR-148b-3p	6.5	2.81	0.7	0.28	<0.0001	9.94
hsa-miR-374a-5p	20.9	5.83	4.2	0.54	<0.0001	10.87
hsa-miR-503-5p*	1.9	3.07	0.2	0.25	<0.0001	12.53
hsa-miR-31-5p*	14.8	9.96	1.8	0.65	<0.0001	15.43
hsa-miR-585	0.7	0.95	0.1	0.06	0.002	15.45

Table 3. *Continued*

miRNA	Carcinoma		Normal		SigGenes *P* value	C/N Fold change
	% Expressing	Mean expression	% Expressing	Mean expression		
hsa-miR-433	0.6	1.52	0.1	0.10	0.006	15.53
hsa-miR-1244	26.7	2.78	4.9	0.17	<0.0001	15.99
hsa-miR-19a-3p*	27.5	5.79	4.0	0.34	<0.0001	17.25
hsa-miR-645	45.6	3.64	15.2	0.21	<0.0001	17.69
hsa-miR-7-5p*	47.2	6.53	11.1	0.30	<0.0001	21.48
hsa-miR-382-5p*	1.3	2.01	0.1	0.09	<0.0001	22.33
hsa-miR-552	9.7	4.12	0.4	0.18	<0.0001	23.27
hsa-miR-17-3p*	19.5	3.15	0.8	0.12	<0.0001	26.19
hsa-miR-409-3p	1.1	3.42	0.1	0.13	0.0001	26.34
hsa-miR-224-5p*	67.3	12.80	10.2	0.46	<0.0001	27.78
hsa-miR-18b-5p*	3.0	2.09	0.2	0.07	<0.0001	30.03
hsa-miR-675-3p	6.7	4.01	0.2	0.11	<0.0001	34.94
hsa-miR-181c-5p*	1.6	4.00	0.1	0.11	<0.0001	35.44
hsa-miR-5701	0.7	1.93	0.1	0.05	0.001	40.11
hsa-miR-455-3p*	18.6	4.58	0.6	0.10	<0.0001	45.10
hsa-miR-18a-5p*	8.7	4.69	0.6	0.10	<0.0001	46.28
hsa-miR-571	1.6	3.94	0.1	0.07	<0.0001	53.53
hsa-miR-135b-5p*	23.2	7.13	1.0	0.12	<0.0001	59.39
hsa-miR-95*	12.1	4.76	0.5	0.07	<0.0001	72.56
hsa-miR-183-5p*	22.4	4.07	0.8	0.05	<0.0001	74.69
hsa-miR-106b-3p*	2.5	1.24	0.1	0.01	<0.0001	97.04
hsa-miR-96-5p*	5.0	4.87	0.1	0.03	<0.0001	166.15
hsa-miR-182-5p	8.6	1.87	0.1	0.01	<0.0001	331.70

[a]miRNA significant only in colon study. [b]miRNA significant only in rectal study.

Table 4. Associations between carcinoma, adenoma and normal tissue among commonly expressed miRNAs (fold change ≤ 0.68 or ≥1.5 and FDR set at 0.05)

	Carcinoma		Adenoma		Normal		C/N		C/A		A/N	
miRNA	% Expressing	Mean expression	% Expressing	Mean expression	% Expressing	Mean expression	SigGenes *P* value	Fold change	SigGenes *P* value	Fold change	SigGenes *P* value	Fold change
Progression of expression (C > AD > N or N > AD > C)												
hsa-let-7i-5p*	97.6	38.62	95.2	21.42	98.3	25.22	0.0002	1.53	<0.0001	1.80	0.001	0.85
hsa-miR-145-5p*	99.7	91.68	100.0	60.51	100.0	187.16	<0.0001	0.49	0.0004	1.52	<0.0001	0.32
hsa-miR-150-5p*	88.3	11.57	93.1	15.30	99.0	28.66	<0.0001	0.40	0.001	0.76	<0.0001	0.53
hsa-miR-17-5p*	99.0	39.75	97.2	22.48	96.2	12.19	<0.0001	3.26	<0.0001	1.77	<0.0001	1.84
hsa-miR-192-5p*	97.9	67.13	99.3	92.66	99.7	99.82	<0.0001	0.67	<0.0001	0.72	0.0005[a]	0.77
hsa-miR-199a-3p	92.8	26.68	69.0	6.03	90.0	13.39	<0.0001	1.99	<0.0001	4.42	<0.0001	0.45
hsa-miR-20a-5p*	97.9	42.78	92.1	23.88	92.1	11.53	<0.0001	3.71	<0.0001	1.79	<0.0001	2.07
hsa-miR-214-3p*	90.4	9.89	51.7	1.82	84.5	4.55	<0.0001	2.17	<0.0001	5.43	<0.0001	0.40
hsa-miR-215	96.2	31.99	98.3	44.45	99.7	50.84	<0.0001	0.63	<0.0001	0.72	0.006	0.87
hsa-miR-21-5p*	99.0	270.61	99.7	148.33	100.0	97.73	<0.0001	2.77	<0.0001	1.82	0.0003	1.52
hsa-miR-223-3p*	98.3	25.77	83.1	11.59	93.9	14.60	0.002[b]	1.76	<0.0001	1.85	0.0008	0.79
hsa-miR-23a-3p*	99.0	119.85	99.3	91.30	100.0	61.13	<0.0001	1.96	0.003	1.31	0.0001	1.49
hsa-miR-24-3p	99.7	74.57	99.7	60.17	100.0	45.87	<0.0001	1.63	0.004	1.24	0.0005	1.31
hsa-miR-25-3p*	95.2	20.32	89.7	12.49	88.0	7.98	<0.0001	2.54	<0.0001	1.63	0.006	1.56
hsa-miR-27a-3p*	97.6	35.25	94.8	26.23	94.8	14.66	<0.0001	2.40	0.0005	1.34	<0.0001	1.79
hsa-miR-29a-3p*	98.3	71.78	98.3	49.46	99.7	32.04	<0.0001	2.24	0.0003	1.45	0.003	1.54
hsa-miR-29b-3p*	92.4	14.56	89.0	8.30	85.2	5.52	<0.0001	2.64	<0.0001	1.75	0.006	1.50
hsa-miR-3651	99.7	50.39	99.7	40.82	100.0	22.11	<0.0001	2.28	0.0001	1.23	<0.0001	1.85
hsa-miR-424-3p*	100.0	37.12	100.0	31.66	100.0	23.70	<0.0001	1.57	0.002	1.17	<0.0001	1.34
hsa-miR-4449	99.3	18.48	100.0	21.29	100.0	13.39	<0.0001	1.38	0.0004	0.87	<0.0001	1.59
hsa-miR-501-3p	96.6	6.60	95.5	5.13	89.7	2.66	<0.0001	2.48	0.007	1.28	<0.0001	1.93
hsa-miR-650	83.5	4.25	97.6	9.60	100.0	15.96	<0.0001	0.27	<0.0001	0.44	<0.0001	0.60
hsa-miR-663b*	100.0	65.27	100.0	48.76	100.0	32.21	<0.0001	2.03	<0.0001	1.34	<0.0001	1.51
hsa-miR-92a-3p*	100.0	82.96	100.0	51.72	100.0	32.27	<0.0001	2.57	<0.0001	1.60	<0.0001	1.60
hsa-miR-93-5p*	96.6	28.90	94.5	20.19	93.5	10.63	<0.0001	2.72	<0.0001	1.43	<0.0001	1.90
Carcinoma–normal and adenoma–normal only (no change from adenoma to carcinoma)												
hsa-miR-10b-5p*	82.8	7.91	75.2	6.52	91.4	11.98	<0.0001	0.66			<0.0001	0.54
hsa-miR-140-3p	82.8	4.83	74.8	3.92	89.0	6.54	0.0007	0.74			<0.0001	0.60
hsa-miR-342-3p*	83.5	8.87	80.3	6.93	96.6	12.96	<0.0001	0.68			<0.0001	0.54
hsa-miR-34a-5p*	95.5	15.65	90.0	13.92	86.9	7.85	<0.0001	1.99			0.0002	1.77
hsa-miR-451a*	74.2	13.34	74.5	18.62	85.9	28.22	<0.0001	0.47			0.003	0.66
hsa-miR-4769-3p	86.9	5.46	84.3	5.04	91.4	6.46	0.02	0.85			0.005[b]	0.69
hsa-miR-6073	85.6	4.12	87.6	3.91	95.9	6.13	<0.0001	0.67			<0.0001	0.64
Carcinoma–normal and carcinoma–adenoma only (no change from adenoma to normal)												
hsa-miR-21-3p*	93.8	15.93	80.3	8.12	81.4	6.66	<0.0001	2.39	<0.0001	1.96		
hsa-miR-222-3p*	96.6	14.97	96.6	9.44	95.5	7.73	<0.0001	1.94	0.0002	1.59		
hsa-miR-331-3p*	94.5	9.85	87.9	5.90	92.4	6.08	<0.0001	1.62	<0.0001	1.67		
hsa-miR-361-5p	89.0	8.74	82.8	6.12	80.8	4.35	<0.0001	2.01	0.0005	1.43		
hsa-miR-375	84.5	14.38	99.0	39.49	99.7	42.48	<0.0001	0.34	<0.0001	0.36		

Table 4. *Continued*

miRNA	Carcinoma		Adenoma		Normal		C/N		C/A		A/N	
	% Expressing	Mean expression	% Expressing	Mean expression	% Expressing	Mean expression	SigGenes *P* value	Fold change	SigGenes *P* value	Fold change	SigGenes *P* value	Fold change
Carcinoma–normal												
hsa-miR-10a-5p*	99.1	26.34			93.9	16.19	0.002[b]	1.63				
hsa-miR-143-3p*	49.1	3.91			74.9	6.23	<0.0001	0.63				
hsa-miR-193b-3p	82.8	6.53			80.4	4.00	0.02	1.63				
hsa-miR-195-5p*	29.6	2.02			78.4	8.03	<0.0001	0.25				
hsa-miR-199a-5p*	85.2	11.92			77.3	5.25	<0.0001	2.27				
hsa-miR-19b-3p	91.4	17.13			78.7	5.86	<0.0001	2.92				
hsa-miR-30c-5p	79.0	4.88			89.0	7.26	<0.0001	0.67				
hsa-miR-425-5p*	80.4	7.46			75.6	4.41	0.002	1.69				
Carcinoma–adenoma only												
hsa-miR-1260b	100.0	125.64	99.7	82.80					<0.0001	1.52		

*miRNA significant only in colon study. [b]miRNA significant only in rectal study.

for those miRNAs more commonly expressed. Examples of miRNAs that were included in this category were hsa-miR-21-5p and hsa-miR-145-5p. Two other categories of dysregulation were those where differences in expression were observed for both the carcinoma and adenoma compared to the normal tissue and those exhibiting differential expression between carcinoma and adenoma and between carcinoma and normal.

Examination of carcinoma/adenoma/normal tissue sequence among those miRNA's expressed in 20–80% of samples showed some slightly different patterns than observed for those miRNAs that were commonly expressed (Table 5; all dysregulated miRNAs for 20 to 80% of participant samples shown in Supplementary Table 5, available at *Carcinogenesis* Online). While for more frequently expressed miRNAS the most common pattern was one of progression of differential expression from normal to adenoma to carcinoma, this was the least common pattern in this group of miRNAs. The most common groups were those that were differentially expressed between carcinoma and normal tissue and between carcinoma and adenoma tissue (i.e. no change from adenoma to normal tissue and those that were only dysregulated between carcinoma and normal tissue).

Comparison of Agilent MicroArray to NanoString platforms

Of the 798 miRNAs on the NanoString platform, 664 had identical miRNA nomenclature with the Agilent Microarray platform. The reliability coefficient for the five repeated samples for the Agilent Platform was 0.98 (See Supplementary Figure 1A, available at *Carcinogenesis* Online). The reliability coefficient for five repeated samples on NanoString was 0.77 (See Supplementary Figure 1B, available at *Carcinogenesis* Online). Evaluation of the distribution of the correlation coefficients for each of the 158 commonly expressed miRNAs between Agilent and NanoString for the 30 samples with measurements on both platforms showed that 46.8% of miRNAs expressed in normal tissue and 74.4% of those expressed in carcinoma tissue had a correlation coefficient between 0.3 and 1.0. Among NanoString replication measurements, correlation coefficients between 0.3 and 1.0 were observed among 84.7% of miRNAs in normal tissue and 60.1% of miRNAs in carcinoma tissue. Among Agilent replication measurements these percentages were 92.1 and 88.8, respectively. A comparison of agreement between Agilent replication samples showed 89% concordant, 4% expression on only one sample and 7% discordant (Figure 1A). Comparable values for NanoString were 78% concordant, 13% expressing in only one sample and 9% discordant (Figure 1B). Comparing NanoString to Agilent for those 664 miRNAs that shared common nomenclature, we observed that 45% were concordant in terms of expression, 43% were expressed on only one of the two platforms and 12% were discordant between platforms (Figure 1C).

Discussion

Our data suggest that many miRNAs are not expressed commonly in colonic tissue, however the majority of miRNAs that are expressed are differentially expressed between carcinoma tissue and normal colonic mucosa when applying an FDR of 0.05. Of those miRNAs expressed in over 80% of the population, 86% were dysregulated; slightly lower numbers of dysregulated miRNA expression between carcinoma and normal tissue was observed for miRNAs expressed less frequently. The majority of miRNAs that are differentially expressed are downregulated in carcinoma tissue. Of those miRNAs that were differentially expressed between carcinoma and normal tissue, roughly half

Table 5. Associations between carcinoma, adenoma and normal tissue among miRNA expressed in 20–80% of subjects (fold change ≤0.68 or ≥1.5 and FDR set at 0.05)

	Carcinoma		Adenoma		Normal		C/N		C/A		A/N	
miRNA	% Expressing	Mean expression	% Expressing	Mean expression	% Expressing	Mean expression	SigGenes *P* value	Fold change	SigGenes *P* value	Fold change	SigGenes *P* value	Fold change
Progression of expression (C > N > AD or N > AD > C)												
hsa-miR-106b-5p*	86.6	9.39	75.2	5.39	62.5	2.74	<0.0001	3.42	<0.0001	1.74	<0.0001	1.96
hsa-miR-146a-5p	79.1	8.39	79.3	12.73	69.6	4.09	0.004[a]	2.05	0.01	0.56	<0.0001	3.21
hsa-miR-20b-5p*	84.5	10.06	70.0	5.10	47.1	1.70	<0.0001	5.92	<0.0001	1.97	<0.0001	3.00
hsa-miR-2117	33.7	1.70	46.2	2.88	63.9	3.67	<0.0001	0.46	0.005	0.59	0.004	0.78
hsa-miR-221-3p*	80.1	8.99	64.8	4.11	48.5	1.90	<0.0001	4.73	<0.0001	2.19	<0.0001	2.16
Carcinoma–normal and adenoma–normal only (no change from adenoma to carcinoma)												
hsa-miR-1291	81.8	4.89	79.7	4.67	73.2	2.82	<0.0001	1.74			<0.0001	1.66
hsa-miR-133b*	11.0	1.09	5.2	0.62	56.7	5.21	<0.0001	0.21			<0.0001	0.12
hsa-miR-142-3p	19.2	1.76	17.2	1.34	37.5	3.27	0.007	0.54			0.0002	0.41
hsa-miR-196b-5p	66.0	10.25	69.0	6.54	62.5	3.52	0.0003	2.91			0.004	1.86
hsa-miR-203a*	66.0	7.77	76.9	9.77	38.1	1.89	<0.0001	4.11			<0.0001	5.16
hsa-miR-30a-5p	18.9	1.14	11.0	0.62	47.8	3.15	<0.0001	0.36			<0.0001	0.20
hsa-miR-378d	2.4	0.16	7.2	0.58	26.1	1.55	<0.0001	0.10			0.0002	0.38
hsa-miR-3976	45.4	2.59	43.8	2.35	34.4	1.21	<0.0001	2.13			0.0005	1.93
hsa-miR-4657	72.5	3.15	73.1	3.34	62.9	2.29	0.003	1.38			0.006	1.46
hsa-miR-497-5p	13.1	0.92	15.5	1.01	59.1	5.10	<0.0001	0.18			<0.0001	0.20
hsa-miR-513c-3p	27.1	2.07	30.7	2.43	46.4	3.45	0.0002	0.60			0.002	0.71
Carcinoma–normal and carcinoma–adenoma only (no change from adenoma to normal)												
hsa-miR-1203	46.7	1.53	60.0	2.26	69.8	2.41	<0.0001	0.64	0.0009	0.68		
hsa-miR-1258	41.6	2.58	49.7	3.92	60.8	3.81	0.005	0.68	0.01	0.66		
hsa-miR-130b-3p*	90.4	7.39	81.4	5.45	77.0	4.22	<0.0001	1.75	0.004	1.36		
hsa-miR-151a-3p*	44.7	3.51	19.3	1.14	21.0	0.64	<0.0001	5.48	<0.0001	3.08		
hsa-miR-15a-5p	51.5	4.36	36.9	2.49	48.1	2.61	0.02	1.67	0.001	1.75		
hsa-miR-192-3p*	13.7	0.87	28.6	2.29	38.8	2.18	<0.0001	0.40	0.001	0.38		
hsa-miR-30e-5p	29.9	1.43	45.2	2.46	54.3	2.61	<0.0001	0.55	0.008	0.58		
hsa-miR-3187-5p	68.7	3.89	47.0	2.64	53.9	2.32	0.01[a]	1.68	0.01[a]	1.47		
hsa-miR-330-3p*	55.3	3.10	71.0	4.76	76.3	5.23	<0.0001	0.59	0.0003	0.65		
hsa-miR-365a-3p*	63.6	6.27	26.2	1.79	41.6	2.55	<0.0001	2.47	<0.0001	3.50		
hsa-miR-3677-3p	62.5	4.31	49.3	3.27	46.0	2.44	<0.0001	1.76	0.003	1.32		
hsa-miR-4251	47.4	3.29	30.4	1.91	35.4	1.56	0.007	2.11	0.007[a]	1.81		
hsa-miR-4421	29.9	1.50	44.0	2.26	46.7	2.40	0.002	0.62	0.007[b]	0.55		
hsa-miR-4469	41.2	1.96	53.4	2.90	60.5	2.82	0.0006	0.70	0.008	0.68		
hsa-miR-5685	24.1	1.47	36.2	2.51	48.5	2.83	<0.0001	0.52	0.01	0.59		
hsa-miR-658	54.3	0.80	70.3	1.23	72.9	1.15	0.0001	0.70	0.005[b]	0.66		
hsa-miR-934	83.5	4.35	39.3	0.72	45.4	0.75	<0.0001	5.81	<0.0001	6.03		
Carcinoma–normal only												
hsa-miR-1207-3p	12.7	0.69			26.1	1.25	0.007	0.55				
hsa-miR-1266	52.2	2.62			35.7	1.38	0.007[a]	1.89				
hsa-miR-196a-5p	61.5	4.69			58.4	2.46	0.004	1.91				
hsa-miR-199b-5p*	29.2	3.16			18.2	0.66	<0.0001	4.78				

Table 5. *Continued*

	Carcinoma		Adenoma		Normal		C/N		C/A		A/N	
miRNA	% Expressing	Mean expression	% Expressing	Mean expression	% Expressing	Mean expression	SigGenes P value	Fold change	SigGenes P value	Fold change	SigGenes P value	Fold change
hsa-miR-3178	33.7	1.22			50.5	1.81	0.0008	0.67				
hsa-miR-3622b-3p	30.6	2.08			13.4	0.53	<0.0001	3.97				
hsa-miR-3972	53.6	3.90			48.5	2.27	0.02	1.71				
hsa-miR-4296	48.5	2.05			38.8	1.15	0.002	1.78				
hsa-miR-4315	0.3	0.04			18.6	1.99	<0.0001	0.02				
hsa-miR-4317	31.6	1.75			22.7	0.99	0.02	1.76				
hsa-miR-4458	65.3	4.09			81.1	5.95	<0.0001	0.69				
hsa-miR-4492	13.7	0.92			29.2	1.49	0.005	0.62				
hsa-miR-5008-3p	27.5	2.45			17.5	0.69	0.0001	3.55				
hsa-miR-532-3p	24.7	2.12			17.2	0.87	0.003	2.45				
hsa-miR-583	89.3	7.07			80.4	4.05	<0.0001	1.75				
hsa-miR-6515-5p	41.2	1.76			85.2	5.17	<0.0001	0.34				
Carcinoma–adenoma only												
hsa-miR-195-3p*	25.9	0.75	16.6	0.37					0.004	2.00		
hsa-miR-324-5p*	26.9	2.79	15.2	1.12					0.01	2.50		
hsa-miR-432-5p*	63.5	2.97	37.4	1.68					0.004[a]	1.77		
Adenoma–normal only												
hsa-miR-148a-3p*			83.8	15.58	75.9	7.65					0.0003	2.04
hsa-miR-195-5p*			35.5	2.05	78.3	8.03					<0.0001	0.26
hsa-miR-28-3p			35.5	0.78	62.8	1.40					<0.0001	0.56
hsa-miR-28-5p			7.9	0.79	26.6	1.62					0.006	0.49
hsa-miR-425-5p*			81.7	7.40	75.5	4.41					0.0009	1.68
hsa-miR-429			63.8	8.85	53.8	4.33					0.0008	2.04
hsa-miR-4312			3.8	0.74	22.1	3.09					<0.0001	0.24
hsa-miR-4785			29.3	1.10	45.5	2.08					0.0006	0.53
hsa-miR-605*			6.6	0.81	26.6	1.64					0.0003	0.50
hsa-miR-671-3p			22.4	0.51	34.1	0.94					0.007	0.54

[a]miRNA significant only in rectal study. [b]miRNA significant only in colon study.

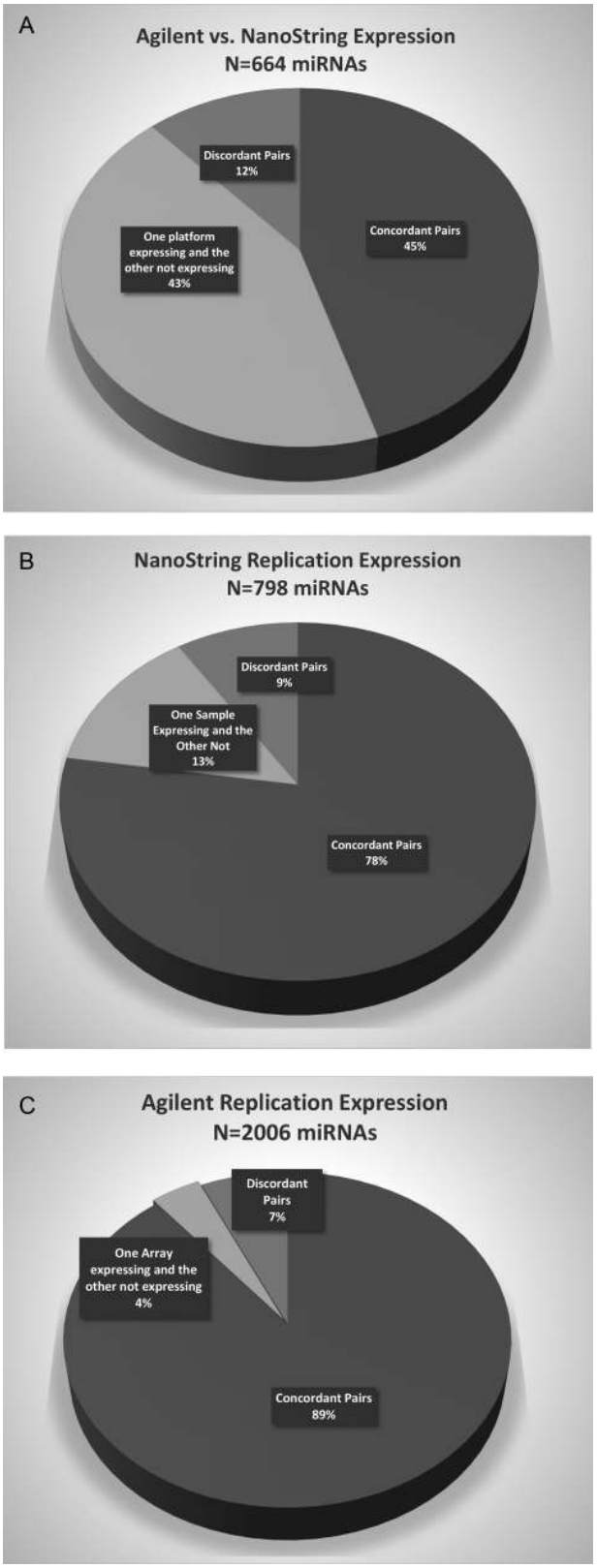

Fig. 1. Agreement between various platforms and repeat samples. (**A**) Agreement of differential expression between Agilent and NanoString platforms. (**B**) Agreement of differential expression for NanoString repeated samples. (**C**) Agreement of differential expression for Agilent repeated samples. Categories of agreement are concordant, discordant and not expressed on one of the compared samples.

followed a pattern of dysregulated miRNA expression that progressed from "normal" to adenoma to carcinoma, while others follow patterns of dysregulated miRNA expression between adenoma and carcinoma or between adenoma and normal tissue or between carcinoma and normal tissue.

Several issues in the interpretation of results need to be addressed prior to comparing differences in expression between various tissue types. One of the first issues is choice of platform and how reproducible results are within platforms. We used the Agilent MicroArray platform because results are highly reproducible, the quantity of RNA needed is readily obtainable from formalin-fixed paraffin-embedded tissue, and it has over 2000 miRNAs available for analysis. These features are important for a study that depends on formalin-fixed tissue and has a discovery component. However, comparing results from different platforms can be challenging given discrepancies in nomenclature, in procedures for handling low levels of expression and background signal, and on normalization methods. In our comparison of Agilent MicroArray and NanoString, 664 miRNAs could be evaluated because of comparable nomenclature. Of those, 313 were not expressed on Agilent Platform and 248 were not expressed on NanoString; only 158 miRNAs that were commonly expressed were available to compare between the two platforms. The reproducibility of NanoString was less than we observed for Agilent. A report by Mestdagh and colleagues (20) examined a set of 16 samples on 12 commercial platforms for miRNA analysis. Of the 196 miRNAs measured on all platforms in that comparison, 66 were expressed on at least two platforms and were the focus of the analysis. They found that Agilent had the highest reproducibility and was one of three platforms that could capture small expression differences.

Consistency in direction of differential expression is important to be able to determine if miRNAs are consistently up- or downregulated. Replication of Agilent microarrays showed 89% agreement, NanoString showed 78% agreement and comparing Agilent to NanoString showed 45% agreement. The major difference in comparison between Agilent and NanoString was a miRNA being expressed by Agilent but not detected on NanoString or vice versa. However, more bothersome are those miRNAs showing differences in direction of expression which account for 11% of the miRNAs shared by Agilent and NanoString, 9% of those in the replication of NanoString and 7% of those in the replication of Agilent. Many of these discrepancies had low levels of expression on one of the replicates. Mestdagh looked at differential expression and found that the average validation of directionality was 54.6% (20). They also assessed truly differentially expressed miRNAs (i.e. called differentially expressed by two different technologies such as PCR and hybridization) agreement rates varied from 35 to 63% (highest being for Agilent and the WaferGen Smart Chip). Of the 66 miRNAs differentially expressed on at least one platform, only two were consistently up or down regulated on all 12 platforms. Overall, we feel that the Agilent platform performed excellent in terms of repeatability and very well when compared to NanoString. However, it should be kept in mind that results can differ by platform used, nomenclature for miRNAs, the number of miRNAs that can be evaluated, and how well they detect low levels of expression.

Normalization is a critical pre-requisite to data analysis given the potential for systematic experimental bias and technical variation that exist for miRNA data (21). Most normalization methods have been developed for mRNA rather than miRNA and should be carefully evaluated when being applied to miRNA data. Unlike mRNA data, miRNA data has numerous miRNAs that are either not expressed or expressed at very

low levels. Median centering, a common method for normalizing mRNA, has little to no impact on miRNA data since the median is around 0. Other methods such as quantile normalization were considered, but after examining raw and normalized data, it appeared to overly reduce expression values at the low end of expression, while having a lesser effect at the upper ends of expression; this would influence calculations such as fold change. We utilized a scaling method to normalize Agilent data, focusing on the upper levels of expression (75th percentile) that we found normalized the variation across arrays while imposing less structure on the data.

The majority (86%) of miRNAs that were expressed commonly in colorectal tissue, were differentially expressed between carcinoma and normal tissue and for both colon and rectal carcinomas when applying an FDR of 0.05. We observed few miRNAs that were not differentially expressed in both colon and rectal tissue. Few studies have evaluated expression profiles for colon and rectal cancer separately. A study of rectal cancer by Gaedcke and colleagues utilized a microarray platform containing 2090 miRNA probes and reported 13 miRNAs that appeared to be rectal cancer specific in that they had not been previously reported in the literature with colon cancer (22). All 13 (miR-492, miR-542–50, miR-584, miR-483-5p, miR-144, miR-2110, miR-652, miR-385, miR-147b, miR-148a, miR-190, miR-26a/b and miR-29c) of these miRNAs were dysregulated in both colon and rectal cancer in our data with the same direction of differential expression as reported by Gaedcke.

The majority of miRNAs were downregulated in both carcinoma and adenomas. Given the number of differentially expressed miRNAs between normal mucosa and carcinoma tissue, our results suggest general dysregulation of miRNAs in carcinomas. This has been suggested by others (23,24). Given the widespread dysregulation of genes observed in carcinoma tissue (25), it is not surprising that widespread dysregulation of miRNAs also exist. While the majority of commonly expressed miRNAs were dysregulated, many had small fold changes; it is not clear what the impact of a 10–30% increase in expression is compared to a twofold increase in miRNA expression. However, fold change as an indicator of importance should be used with caution (26). This is especially true when looking at those miRNAs less commonly expressed, where the majority of samples may have minimal or no expression and there is increase or decrease in carcinoma miRNA expression resulting in a large fold change but small absolute change in expression. It is important to broaden our understanding of the implications of less frequently expressed miRNAs and the impact of low levels of expression. We examined frequency of expression in terms of carcinoma molecular phenotype and did not see a pattern of infrequent expression specific to any carcinoma phenotype. Further examination of differential expression by tumor molecular phenotype is needed, but beyond the scope of this manuscript.

Others have suggested that there is a progression in miRNA expression as tissue goes from 'normal' to adenoma to carcinoma (27). We observed that roughly half of the miRNAs that were differentially expressed between carcinoma and normal tissue also were differentially expressed between adenoma and normal and adenoma and carcinoma, representing a progression in miRNA expression. However, other patterns suggest that some miRNAs are altered at the normal to adenoma stage, or at the adenoma to carcinoma stage. Some miRNAs were only differentially expressed between carcinoma and normal, suggesting that the difference between normal and adenoma was not of the same magnitude.

Our study has several strengths. First, we have a large sample of paired carcinoma and normal tissue; additionally we have paired triplicate samples (carcinoma, normal, adenoma) on 290 people. This large sample has enabled us to perform a validate findings in our population by splitting the sample into two separate groups for analysis and only designating those miRNA as being dysregulated when they were significantly differentially expressed in both. Our large sample also has allowed us to gain insight into the patterns and distribution of miRNA expression. Because of the large sample, we were able to document that hundreds of miRNAs are only expressed in smaller subsets of the population. Likewise, our sample will enable us to evaluate these differences in more detail to hopefully determine factors that influence their expression. Our inclusion of carcinomas, adenomas, and normal tissue allows us to identify miRNAs associated with disease progression as well as miRNAs that only become dysregulated at specific points in the carcinogenic process. A major strength of our study is the paired 'normal' tissue that enables us to better quantify dysregulation. Any miRNA expression that is altered by genetic or lifestyle factors will be automatically controlled using this study design. We also consider using the Agilent platform a study strength. This has enabled us to evaluate over 2000 miRNAs that will enable discovery of new important miRNAs as well as validate previously identified miRNAs. The Agilent Microarray Platform was extremely reliable with reasonable agreement with the NanoString platform.

In conclusion, we have shown that global miRNA dysregulation occurs in colorectal carcinomas and that few differences occurred by carcinoma subsite. Given the global dysregulation, it will be important to identify smaller subsets of key miRNA predictors that can be used to determine disease progression from normal to adenoma to carcinoma as well as key miRNAs that are associated with disease development and survival. Given the scope of dysregulated miRNAs, the influence of total dysregulated miRNA burden rather than individual miRNAs needs consideration. These results suggest the importance of miRNAs in colorectal cancer, however to fully understand the significance of miRNAs to the carcinogenic process will involve better characterization of those miRNAs commonly expressed in colorectal tissue as well as those expressed more sporadically.

Supplementary material

Supplementary Tables 1–5 and Figure 1 can be found at http://carcin.oxfordjournals.org/

Funding

National Cancer Institute at the National Institutes of Health (CA163683, CA48998 and CA61757).

Acknowledgements

The contents of this manuscript are solely the responsibility of the authors and do not necessarily represent the official view of the National Cancer Institute. We acknowledge Sandra Edwards for data oversight and Dr. Bette Caan and the staff at the KPMCP for their contribution to data collection and to Dr. Chris Corcoran at Utah State University for his input into the statistical analysis. M.L.S. obtained funding, oversaw study and wrote the manuscript. J.S.H. conducted statistical analysis. D.F.P. gave input into the statistical analysis. J.R.S. provide overall guidance to statistical analysis. W.S. reviewed slides for identification of carcinoma, adenoma and normal tissue. L.E.M. coordinated study data and

assisted with data interpretation. E.W. and M.H. conducted miRNA analysis. R.K.W. provided oversight to laboratory analysis. All authors read and approved the manuscript.
Conflict of Interest Statement: None declared.

References

1. Iorio, M.V. et al. (2009) MicroRNAs in cancer: small molecules with a huge impact. J. Clin. Oncol., 27, 5848–5856.
2. Ambros, V. (2004) The functions of animal microRNAs. Nature, 431, 350–355.
3. Murray, B.S. et al. (2010) An in silico analysis of microRNAs: mining the miRNAome. Mol. Biosyst., 6, 1853–1862.
4. Arora, S. et al. (2013) miRNA-transcription factor interactions: a combinatorial regulation of gene expression. Mol. Genet. Genomics, 288, 77–87.
5. Gartel, A.L. et al. (2008) miRNAs: Little known mediators of oncogenesis. Semin. Cancer Biol., 18, 103–110.
6. Nam, S. et al. (2009) MicroRNA and mRNA integrated analysis (MMIA): a web tool for examining biological functions of microRNA expression. Nucleic Acids Res., 37, W356–62.
7. Drusco, A. et al. (2014) MicroRNA profiles discriminate among colon cancer metastasis. PLoS One, 9, e96670.
8. Michael, M.Z. et al. (2003) Reduced accumulation of specific microRNAs in colorectal neoplasia. Mol. Cancer Res., 1, 882–891.
9. Volinia, S. et al. (2006) A microRNA expression signature of human solid carcinomas defines cancer gene targets. Proc. Natl. Acad. Sci. USA, 103, 2257–61.
10. Yang, L. et al. (2009) MicroRNA and colorectal cancer. World J. Surg., 33, 638–646.
11. Acunzo, M. et al. (2015) MicroRNA and cancer–a brief overview. Adv. Biol. Regul., 57, 1–9.
12. Slattery, M.L. et al. (1997) Energy balance and colon cancer–beyond physical activity. Cancer Res., 57, 75–80.
13. Slattery, M.L. et al. (2003) Energy balance and rectal cancer: an evaluation of energy intake, energy expenditure, and body mass index. Nutr. Cancer, 46, 166–171.
14. Slattery, M.L. et al. (2000) Use of archival tissue in epidemiologic studies: collection procedures and assessment of potential sources of bias. Mutat. Res., 432, 7–14.
15. Griffiths-Jones, S. (2006) miRBase: the microRNA sequence database. Methods Mol. Biol., 342, 129–138.
16. Suyundikov, A. et al. (2015) Accounting for dependence induced by weighted KNN imputation in paired samples, motivated by a colorectal cancer study. PLoS One, 10, e0119876.
17. Agilent Technologies Inc. (2013) Agilent GeneSpring User Manual. Santa Clara, CA.
18. Tusher, V.G. et al. (2001) Significance analysis of microarrays applied to the ionizing radiation response. Proc. Natl. Acad. Sci. USA, 98, 5116–5121.
19. Benjamini, Y. et al. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J. Roy. Stat. Soc., 57, 289–300.
20. Mestdagh, P. et al. (2014) Evaluation of quantitative miRNA expression platforms in the microRNA quality control (miRQC) study. Nat. Methods, 11, 809–815.
21. Meyer, S.U. et al. (2010) Normalization strategies for microRNA profiling experiments: a 'normal' way to a hidden layer of complexity? Biotechnol. Lett., 32, 1777–1788.
22. Gaedcke, J. et al. (2012) The rectal cancer microRNAome–microRNA expression in rectal cancer and matched normal mucosa. Clin. Cancer Res., 18, 4919–4930.
23. Iorio, M.V. et al. (2012) Causes and consequences of microRNA dysregulation. Cancer J., 18, 215–222.
24. Deng, S. et al. (2008) Mechanisms of microRNA deregulation in human cancer. Cell Cycle, 7, 2643–2646.
25. Slattery, M.L. et al. (2015) Gene expression in colon cancer: A focus on carcinoma site and molecular phenotype. Genes Chromosomes Cancer, 54, 527–41.
26. Wang, B. et al. (2013) Challenges for MicroRNA microarray data analysis. Microarrays (Basel), 2, 34–50.
27. Oberg, A.L. et al. (2011) miRNA expression in colon polyps provides evidence for a multihit model of colon cancer. PLoS One, 6, e20465.