ORIGINAL ARTICLE—ALIMENTARY TRACT

Differential expression of miR-144* as a novel fecal-based diagnostic marker for colorectal cancer

Murugan Kalimutho · Giovanna Del Vecchio Blanco ·
Serena Di Cecilia · Pierpaolo Sileri · Micaela Cretella ·
Francesco Pallone · Giorgio Federici · Sergio Bernardini

Abstract

Background MicroRNAs (miRNA) are tiny, noncoding, small, endogenous RNAs that play major roles in neoplastic transformation and could therefore offer a better quantitative and noninvasive method for the diagnosis and prognosis of colorectal cancer (CRC) using feces. In the present study, we screened feces for 648 miRNAs and analyzed the role of miR-144* as a potential CRC diagnostic marker.

Methods Fecal miRNA expression was profiled with RT-pre-amplification-qPCR, and the stability was determined using both endogenous and exogenous miRNA by RT-qPCR. ROC analysis was performed to enhance the diagnosing power of the CRC patients' fecal specimens.

Results We detected 39% of all the miRNAs screened in feces. Endogenous miRNAs are more stable over time and temperature, while exogenous miRNAs degraded rapidly. miR-144* was overexpressed in feces, suggesting that it could be a potent candidate diagnostic marker for CRC detection, with a sensitivity of 74% and a specificity of 87% ($n = 75$, $p < 0.0001$). Moreover, RT-qPCR analysis showed that miR-144* was also overexpressed in paired CRC tissues, thus suggesting its possible utilization as a diagnostic marker.

Conclusions We demonstrated that miRNAs are stable in the fecal microenvironment, and that, among them, miR-144* represents a novel fecal-based diagnostic marker for CRC screening. Nevertheless, our data need to be validated in a large cohort of subjects.

Keywords miR-144* · Fecal miRNAs · CRC · miR-378

Abbreviations
miRNA MicroRNA
fDNA Fecal DNA
CRC Colorectal cancer

M. Kalimutho · S. Di Cecilia · F. Pallone · G. Federici ·
S. Bernardini (✉)
Department of Internal Medicine, University of Rome
"Tor Vergata", Via Montpellier 1, 00133 Rome, Italy
e-mail: bernardini@med.uniroma2.it

Present Address:
M. Kalimutho (✉)
Centre for Cancer Research and Cell Biology (CCRCB),
Queen's University, Belfast, Northern Ireland, UK
e-mail: mceric_rugan@hotmail.com

G. Del Vecchio Blanco · P. Sileri · M. Cretella · F. Pallone
Gastroenterology Unit, University Hospital Tor Vergata,
University of Rome Tor Vergata, Rome, Italy

P. Sileri
Surgical Unit, University Hospital Tor Vergata,
University of Rome Tor Vergata, Rome, Italy

Introduction

MicroRNAs (miRNAs) are a small class of regulatory RNA molecules that modulate the activity of specific mRNA targets and display important roles in a wide range of physiological and pathological processes [1–4]. They guide the RNA-induced silencing complex to target sites, which are thought to be most prevalent in the 3′ untranslated region of some mRNAs [5]. Moreover, they also potentially target up to one-third of all human coding genes, making their role in cellular biology even more important [6]. Initially, miRNAs were discovered in *Caenorhabditis elegans* as small temporal RNAs (stRNAs) that regulate developmental transitions [7]. Since then,

hundreds of members of the class of miRNAs have been found in worms, flies, plants and mammals. Although the mechanisms through which miRNAs regulate their target genes are still gradually being discovered, the finding that at least some miRNAs feed into the RNA interference (RNAi) pathway has provided a starting point for the discovery and elucidation of the biological roles of miRNAs in normal physiology and human diseases, particularly cancers (for more on this topic, refer to [7, 8]).

At the time that this manuscript was written, 1413 human miRNAs had been reported bioinformatically in miRBase (http://www.mirbase.org/cgi-bin/query.pl?terms=hsa-miR), and these require experimental validation. These molecules are implicated in numerous cellular processes, including development, differentiation, proliferation, apoptosis, DNA damage repair and stress responses, and they are frequently deregulated in cancers [9, 10]. miRNA expression varies with tumor type, including colorectal cancer (CRC) [1, 2, 9–12]. Furthermore, miRNA expression signatures have been shown to be clustered according to tumor type, and to predict prognosis [11–14].

Recent reports have proposed the possibility of detecting miRNAs from serum and blood-based bodily fluids as novel targets to interpret disease state and as biomarkers in various malignancies [15–17]. miRNA expression is tissue specific, and it has unusually high stability in formalin-fixed tissues [18–20] and in bodily fluids [15–23]. In view of this observation, and reports published previously, the expression profile of deregulated miRNAs may represent an enhanced quantitative and diagnostic tool for cancer detection. Although various molecular methods have recently been proposed for CRC detection, most of them have presented relatively low sensitivities as screening modalities.

Cells exfoliated from malignant colonocytes undergo genetic changes that confer upon them the ability to escape from programmed cell death and shed more frequently than normal colonocytes in feces [24, 25]. This phenomenon allows us to distinguish the altered expressions of deregulated miRNAs in order to detect CRC using feces (for a detailed explanation, see [24–27]).

In this study, we assessed whether the altered expression of miRNAs detected in feces could serve as a good candidate biomarker for CRC detection. We profiled 648 miRNAs in pooled samples from subsets of diseased and healthy individuals and determined whether miRNAs are stably detected in feces. Moreover, RT-qPCR was performed to determine whether differential miRNA expression can be robustly and reliably measured in feces. With regards to the generated microarray data, we validated a chosen set of miRNAs that could diagnose CRC accurately with acceptable sensitivity and specificity. We showed that endogenous miRNAs are stably detected in feces, and that,

among them, miR-144* represents a novel fecal-based diagnostic marker for CRC screening.

Materials and methods

Feces and tissues samples

Fecal specimens were collected from patients who underwent colonoscopy following a standardized procedure, as described previously [26, 27]. Briefly, patients were instructed to collect an aliquot of feces in a sealed plastic container and transport it to the hospital in an ice bag. The samples were stored at $-20°C$ immediately upon receipt with fecal stabilization buffer. The samples were then processed according to the standard protocol, as described previously [26]. During colonoscopy, the colon was visualized until the cecum in all patients, independent of disease state, and the sizes and locations of any lesions were recorded. Later, the diagnosis of cancer was confirmed histologically and pathological stages were scored. Surgical tissues were also collected from another set of cancer patients who underwent colonoscopy and surgery in our institute and kept at $-80°C$. All subjects gave their informed consent prior to participation in this study. The study was approved by our institutional review board. miRNA analysis was conducted blindly by two technicians who evaluated the pooled and RT-qPCR assays.

Isolation of miRNAs from the feces of CRC and healthy subjects

Total miRNA was isolated from approximately 20–30 mg of feces of CRC and healthy individuals ($n = 75$) using the mirVana^{TM} miRNA isolation kit with minor modifications from the manufacturer's (Ambion, Inc.) guidelines. For the initial lysis step, we used 800 μl of lysis/binding solution diluted in feces and vortexed vigorously. Later, the mixture was subjected to centrifugation at 13000 rpm for 1 min. The supernatant was collected and processed according to the instructions that came with the mirVana isolation kit. The concentrations of all miRNA samples were quantified using NanoDrop 1000 (Nanodrop, Wilmington, DE, USA), and an equivalent amount of the total miRNAs of each patient was used for RT-qPCR analysis.

Fecal miRNA stability assay

Feces samples were spiked with synthetic cel-miR-238 from *C. elegans*, which has no sequence homology to human miRNAs, as described by Mitchell et al. [16] (miRBase accession number: MIMAT000093-UUUGU ACUCCGAUGCCAUUCAGA; Primm SRL, Milan, Italy).

Molar concentrations to be spiked in were derived empirically to produce C_q values that were comparable to those of moderately abundant endogenous miRNAs measured in fecal specimens. Briefly, 20–30 mg of the fecal sample containing 2 µl of 10 µM cel-miR-238 and the negative control were re-suspended in duplicate in 400 µl of 2X denaturing solution (Ambion), which contains RNase inhibitor, or in PBS without RNase inhibitor. Samples were then processed immediately with the mir-Vana lysis/binding solution. An equivalent amount of total miRNA from each sample was reverse transcribed, and the level of cel-miR-238 was quantified based on RT-qPCR. For the temperature stability analysis, equal amounts of each fecal sample were placed at 4 and 25°C, and the levels of miR-16 and cel-miR-238 were monitored over periods of 0, 2, 4 and 6 days and quantified by standard curve analysis.

RDML data guidelines

Raw miRNA expression, experimental annotation, and sample annotation are available in the RDML data format [28].

Microarray miRNA reverse transcription

Feces from 15 CRC patients and 15 healthy controls were pooled to analyze the expression of 648 miRNAs. Total miRNA was extracted using the mirVana^TM miRNA isolation kit, and then 8 µl of the total RNA (50 ng/µl) (the same amount of total miRNA was pooled for each patient, quantified by a NanoDrop 1000) were reverse transcribed using an miRNA reverse transcription kit (Applied Biosystems) in combination with the stem-loop Megaplex primer pool (Applied Biosystems), allowing the simultaneous reverse transcription of 648 miRNAs and 19 endogenous controls [23, 28]. Briefly, 8 µl of the total RNA (50 ng/µl) were supplemented with RT primer mix (10×), RT buffer (10×), MultiScribe reverse transcriptase (10 U/µl), dNTPs with dTTP (0.5 mM each), MgCl$_2$ (3 mM), and AB RNase inhibitor (0.25 U/µl) in a total reaction volume of 80 µl. For RT reactions with subsequent pre-amplification, the reaction volume was proportionally reduced to 5 µl. The concentration of each stem-loop primer in the RT reaction mix was 1 nM, a 50-fold dilution compared with a singleplex RT reaction, ensuring minimal nonspecific interactions between the different stem-loop primers. To increase the reverse transcription efficiency, a pulsed RT reaction was used (40 cycles of 16°C for 2 min, 42°C for 1 min and 50°C for 1 s, followed by a final reverse transcriptase inactivation at 85°C for 5 min).

Pre-amplification of miRNA cDNA

Megaplex RT product (5 µl) was pre-amplified using Applied Biosystems' TaqMan PreAmp Master Mix (2×) and the PreAmp Primer Mix (5×) in a 25-µl PCR reaction. The primer pool consisted of forward primers (50 nM) that were specific for each of the 667 miRNAs, including 19 endogenous control small RNAs and a universal reverse primer (50 nM) (Applied Biosystems). The pre-amplification cycling conditions were as follows: 95°C for 10 min, 55°C for 2 min and 75°C for 2 min, followed by 14 cycles of 95°C for 15 s and 60°C for 4 min [29, 30].

Real-time qPCR

For each cDNA sample, 648 miRNAs and 19 endogenous control small RNAs were profiled using a gene maximization PCR plate setup in a 384-well plate. This approach allowed us to profile one sample per 384-well plate. The RT product was diluted 400-fold; when pre-amplification was applied, the dilution factor was 1600. PCR amplification reactions were carried out in a total volume of 8 µl, containing 4 µl of TaqMan Master Mix (Applied Biosystems), 1 µl of cDNA and 3 µl of miRNA TaqMan probe and primers (primer details can be accessed through the Applied Biosystems website: http://www.appliedbiosystems.com). Cycling conditions were as follows: 95°C for 10 min, followed by 40 cycles of 95°C for 15 s, and 60°C for 1 min. All PCR reactions were performed on the 7900HT RT-qPCR system (Applied Biosystems). Raw C_q values were calculated using SDS software v.2.1 with automatic baseline settings and a threshold of 0.224. The crossing point between the baseline-corrected amplification curve and the threshold line is called the quantification cycle (C_q) (according to RDML guidelines, http://www.rdml.org) [28].

Selection of miRNAs/small RNA controls that resemble the mean expression value

In order to perform the robust and unbiased selection of genes with expression levels that had the highest correlations with the mean expression levels, we used the geometric mean calculation method [30, 31]. In brief, for each miRNA and small RNA control, we initially calculated the geometric mean of the RT-pre-amplification-q-PCR microarray data, and then we normalized our miRNA expression with the calculated geometric mean. Next, the standard deviations of the C_q values of both groups were determined for every miRNA and small RNA control. The miRNAs and/or small RNA controls with the lowest standard deviations were considered to be good candidate reference genes for RT-qPCR analysis [30].

RT-qPCR analysis of miR-144*, miR-532-3p, and miR-378

We analyzed feces from 35 CRC patients and 40 healthy individuals for the expression of miR-144*, miR-532-3p, and miR-378 using multiplex RT-qPCR. After RNA extraction, the samples were retrotranscribed using a standard ABI RT reaction that contained 4 μl RNA (200 ng), 0.40 μl deoxynucleotide triphosphate mix, 3.0 μl 50 U/μl MultiScribe reverse transcriptase, 2.0 μl 10× RT buffer, 0.25 μl AB RNase inhibitor, 7.35 μl water, and 1.0 μl of each of the three primers in a total reaction mixture of 20.0 μl. The RT reaction was carried out at 16°C for 30 min and 42°C for 30 min. The reverse transcription products were diluted onefold with water (final concentration of 5 ng/μl), and 3.0 μl of cDNA were used for RT-qPCR assays. RT products were processed as follows: 95°C for 10 min, and then 95°C for 15 s and 60°C for 30 s for up to 40 cycles. The experiments were repeated twice with three technical replicates, and then the geometric mean of C_q for each sample was calculated for gene expression via the $2^{-\Delta\Delta C_q}$ method.

Isolation and quantification of miR-144* from CRC paired surgical tissues

Fifteen CRC paired surgical tissues ($n = 30$) were used for miR-144* expression analysis. Briefly, total RNA was extracted from tissues using the miRNeasy mini kit (Qiagen, Hilden, Germany), following the manufacturer's instructions. The final elution volume was 30 μl. The concentrations of all miRNA samples were quantified using the NanoDrop 1000. RT-qPCR detection was carried out for each patient sample using the method described above for fecal miRNA detection.

Statistical analysis

Data analysis was performed using the MedCalc (v.9.2.0.1) software package. Receiver operating characteristic (ROC) curve analysis was used to determine the best cutoff value for the highest assay sensitivity and specificity. In addition, contingency table analysis, the independent paired Student's t test, and one-way ANOVA were performed to evaluate the significance of the data where appropriate. Two-sided P values ≤0.05 were considered significant.

Results

Occurrence of miRNAs in feces

We first determined whether miRNAs could be detected in feces, which contain inhibitors and materials of nonhuman origin, such as DNA, RNA, proteins, polysaccharides and possible noncoding RNAs [26]. We isolated miRNAs from the feces of healthy and CRC individuals and checked the expression levels of miR-16 and miR-21 in the first instance. As expected, we detected the expression of these miRNAs in feces, independent of their disease state (data not shown).

Endogenous miRNAs are stably detected in the fecal microenvironment

Before checking for candidate miRNAs that can discriminate healthy from diseased subjects, we first carried out an experiment to verify the stability of miRNAs in feces. We assumed that endogenous fecal miRNAs are protected from degradation through their association with the macromolecules in feces. To test our assumption, we performed a spiked experiment with a known nonhuman miRNA sequence and measured the degradation pattern caused by endogenous RNase activity. We used cel-miR-238 as a model of exogenous miRNA, as this lacks sequence homology with human miRNAs. We were able to amplify the spiked cel-miR-238 in RT-qPCR analysis after inhibiting the existing RNase activity (Fig. 1A-a). A control experiment was carried out where the RNase activity was not inhibited, and it was found that cel-miR-238 was rapidly degraded within 10 min of exposure to the fecal samples (Fig. 1A-b). Conversely, when RNase inhibitor was added, we observed no decrease in the cel-miR-238 level. In addition to the spiking experiment, we also analyzed the expression of miR-16 in our spiked samples, and found that the miR-144* level was unaltered (data not shown). Thus, we demonstrated that endogenous miRNAs are more stable and resistant to RNases than exogenous miRNA in feces (Fig. 1a).

Fecal miRNAs are time and temperature stable

To further investigate the pre-analytical steps involved in fecal collection and transportation, we analyzed miR-16 stability profiles in samples stored at two different temperatures (4°C and 25°C) over a period of six days. We found that feces kept at 4°C (Fig. 1b, d) are more stable than feces stored at room temperature (Fig. 1c, d). We detected a slight degradation pattern of endogenous miR-16 in feces stored at room temperature during the fourth and sixth days relative to 4°C (Fig. 1d). Thus, these data indicate that endogenous miRNAs are relatively time and temperature stable in feces.

The presence of miRNAs in feces and the selection of a putative candidate biomarker

Having demonstrated that miRNAs are detectable and stable in feces, we further evaluated the deregulated expression of
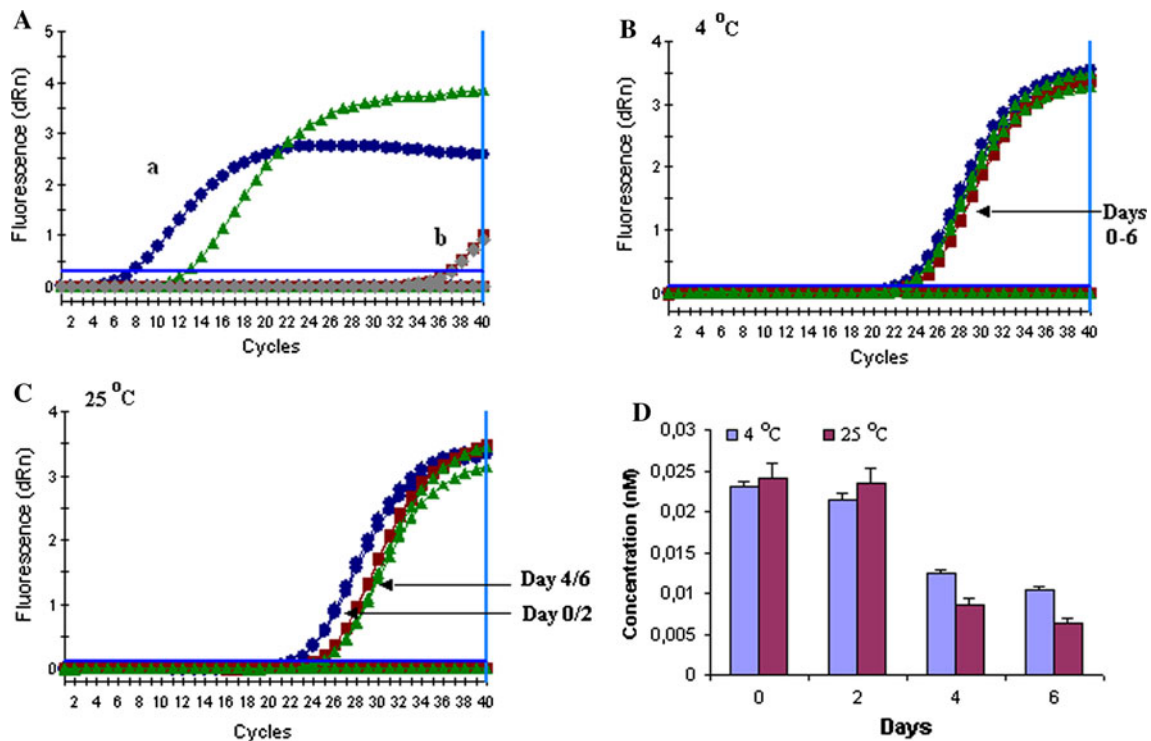
Fig. 1 Analysis of the stability of exogenous cel-miR-238 and endogenous miR-16 in feces over time and at different temperatures. miRNAs were assessed by an RT-qPCR assay. **a** A Synthetic version of the *Caenorhabditis elegans* miRNA (cel-miR-238) was spiked into feces after the addition of an equal volume of 2× denaturing solution (to inhibit RNases). The feces were then incubated at room temperature for 10 min and a very high amplification was detected (curve *a*). The same amount of cel-miR-238 was spiked into feces prior to the addition of an equal volume of 2× denaturing solution, and a very low signal was detected at end of the RT-qPCR cycle (curve *b*). miR-16 stability analysis at 4°C (**b**) and 25°C (**c**) over a period of six days. No changes were observed when the feces were incubated at 4°C, and a slower degradation pattern was detected at 25°C on the fourth and sixth days. **d** Histogram demonstrating minor changes in the expression of miR-16 over a similar period of incubation (4–6 days) to that shown in **b** and **c**, using standard curve analysis

miRNAs in a subset of CRC and healthy individuals ($n = 30$). To do this, we employed an RT-preamp-qPCR microarray technique. A total of 648 miRNAs and 19 endogenous control small RNAs were analyzed, and we observed that 39% (253/648) of the miRNAs in feces. Out of this 39%, 82% (206/252) of the miRNAs were detected in healthy individuals, while 93% (234/252) were detected in CRC subjects. Next, we constructed a linear regression plot to determine the differences in CRC and healthy individuals based on the RT-pre-amp-qPCR microarray data. We found that at least 20% of the miRNAs showed different expression levels in the feces from CRC and healthy individuals ($R^2 = 0.80$) (Fig. 2a). Following the identification of differentially expressed miRNAs in feces, we identified 21 miRNAs that exhibited greater than fivefold upregulating in the CRC samples, although variations in the samples from individual patients may have influenced this result (Table 1; Fig. 2). We then narrowed down the miRNAs by ordering them from the highest expression levels to the lowest (Table 1). Based on this, miR-144* and miR-532-3p showed the greatest increases in their expression levels in CRC. We then re-evaluated the scatter plot to determine which of the miRNAs could serve as a normalizer, since its expression was not influenced by disease state (Fig. 2b). Thus, the miRNAs that were at the boundary of the scatter plot were possible candidates. In the end, we chose miR-378, which had the lowest standard deviation of C_q in the RT-preamp-qPCR microarray experiment (SD; 0.00026) (Fig. 2b).

Furthermore, the pooled analysis revealed several miR-NAs with differential expression, in accordance with literature observations (Table 1; Fig. 2b). However, we could not consider such miRNAs as candidates, as insignificant data were obtained for them from our preliminary investigation (data not shown). We also avoided selecting endogenous control small RNAs for two main reasons: (1) we could only detect three endogenous control small RNAs (RNU24, RNU44 and U6) out of the 19 endogenous controls analyzed, and these had inconsistent expression levels between CRC and healthy individuals (data not shown); (2) the data for the endogenous controls had relatively high standard deviations and the highest fold changes in expression levels observed in the RT-preamp-qPCR analysis.
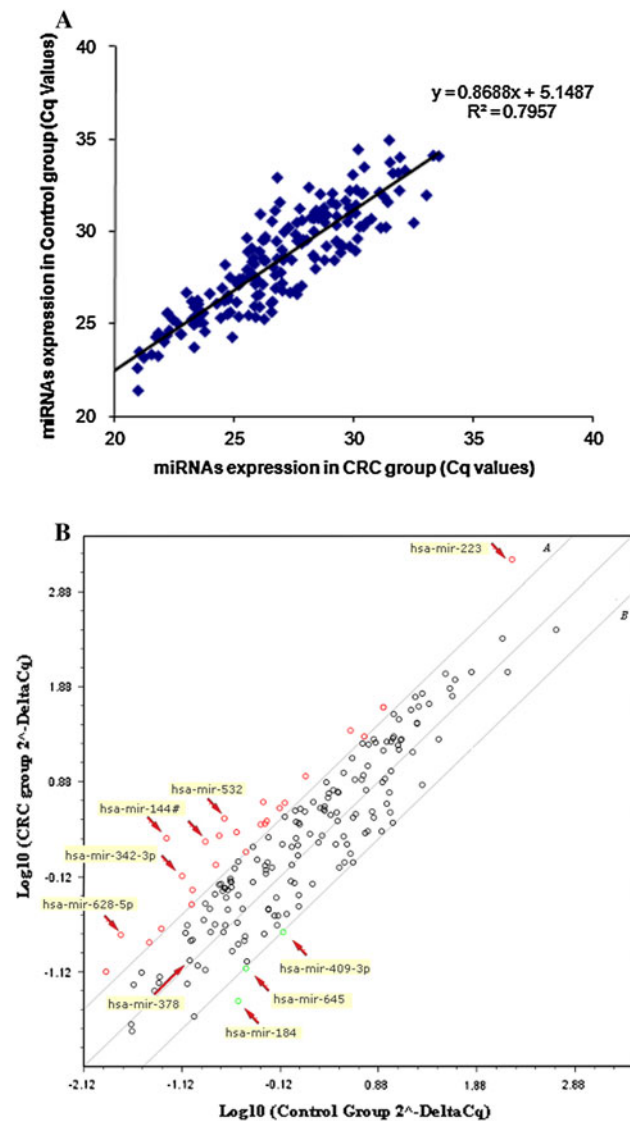
Fig. 2 Regression plot analysis and relative fold changes in expression levels of miRNAs in feces of pooled samples from healthy and CRC subjects. **a** Linear regression analysis indicated that at least 20% of the miRNAs showed different levels of expression in the feces of CRC and healthy subjects. **b** A total of 648 miRNAs were profiled with the RT-pre-amplification-qPCR miRNA microarray. The scatter plot represents the relative fold changes in expression levels between control and CRC groups. The *arrows* show upregulated and downregulated miRNAs. The *circles above line A* show upregulated miRNAs, and the *circles below line B* show downregulated miRNAs, while the *circles in between line A and B* represent nonsignificantly expressed miRNAs. The *middle line* of the graph represents the boundary between upregulated and downregulated miRNAs

Correlation of miR-378 normalization in the feces
and tumor samples

The linear regression analysis showed that miR-378 had the lowest R^2 value with miR-144* in both healthy and CRC subjects (Fig. 3a), whereas miR-378 had a slightly higher correlation in CRC with miR-532-3p (Fig. 3b), where 60%

of the variance was shared between miR-378 and miR-532-3p. In contrast, only 31% of the variance was shared in the healthy subjects. Then we analyzed whether both of these miRNAs were correlated with our selected normalizer. We found that miR-532-3p was positively correlated ($R^2 = 0.53$) (Fig. 3c), while miR-144* which had the lowest shared variance ($R^2 = 0.15$). These data show that the expression of miR-144* is independent of the level of miR-378 in the fecal sample, as only 18% (for the healthy subjects) and 31% (for CRC subjects) of the variance was shared (Fig. 3a). We then analyzed the relationship between miR-378 and miR-144* in tissue samples, and found that miR-144* also had the lowest shared variance, confirming what we observed in fecal specimens (Fig. 3d).

miR-144* as a novel diagnostic marker
for CRC detection

Finally, we evaluated the potential use of our selected miRNAs as a noninvasive fecal-based screening marker for CRC. Based on the C_q values obtained, which were normalized to miR-378 [geometric mean expression of all patients ($n = 75$): 31.97 with SD of 1.75, Fig. 4] in a subset of fecal samples ($n = 35$ CRC and $n = 40$ healthy subjects, Table 2), we detected a significant alteration of ΔC_q between these two groups, as shown in Fig. 5 (together with the box and whisker plot for miR-144*, Fig. 4). Based on this, we detected CRC with a sensitivity of 74.29% (95% CI 56.7–87.5) and a specificity of 87.20% (95% CI 72.6–95.70), AUC ROC of 0.829 (Fig. 5, left panel). Nevertheless, we could only detect 63% of the CRC cases at a specificity of 95%. Next, the cumulative fold change in miR-144* for the whole group showed an 11-fold upregulation for CRC compared to healthy individuals (Table 2). Taken together, the RT-preamp-qPCR microarray and RT-qPCR data led us to conclude that miR-144* is a novel candidate diagnostic marker for CRC screening. Apart from this, the elevated expression of miR-532-3p could not be used to distinguish CRC patients from healthy individuals, as shown in the box and whisker plot, $p = 0.8676$ (Figs. 4, 5, right panel). Moreover, the expression of miR-532-3p was positively correlated with that of miR-378 (Fig. 3b, c).

Upregulation of miR-144* expression in CRC tissues

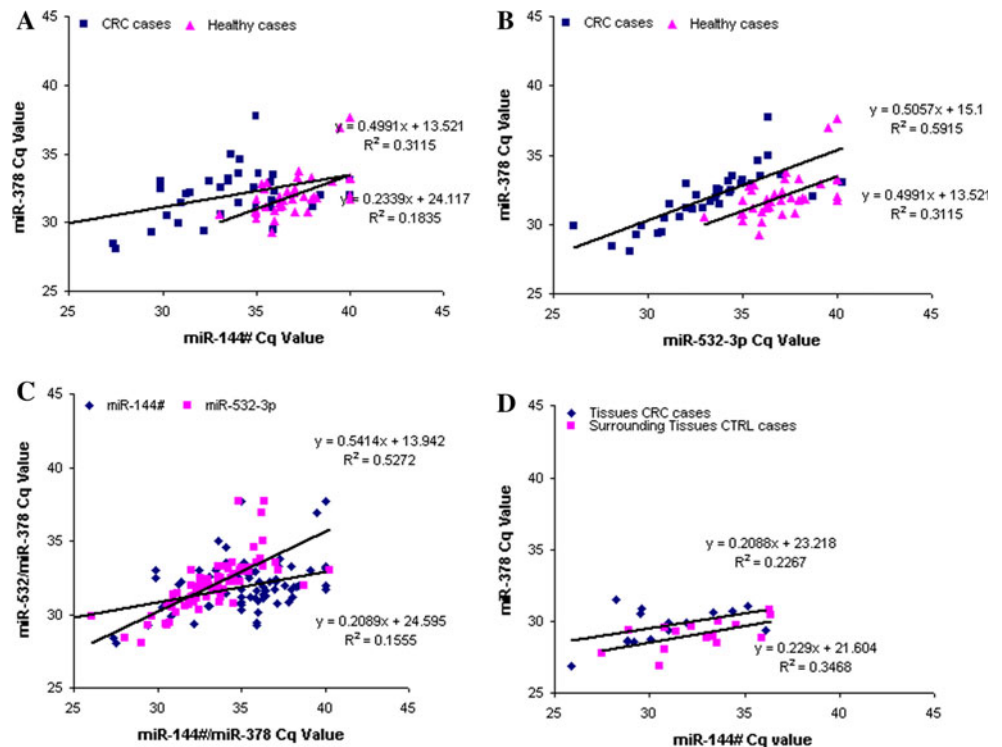In order to verify that the levels of miR-144* detected in feces are similar to those in tissue samples, we analyzed 15 CRC paired surgical surrounding tissues by RT-qPCR. As expected, a strong upregulation of miR-144* expression in CRC tissue samples (~4.99-fold) was observed compared to the expression in the surrounding tissues (Table 3). However, three patients showed downregulation of miR-144*.

Table 1 miRNAs that showed at least fivefold changes in their expression levels in feces of colorectal cancer patients compared to feces of healthy individuals

MicroRNA	MirBase no.	Raw C_q of healthy	Raw C_q of CRC	Fold change
hsa-mir-886-5p	MIMAT0004905	29.57	26.25	5.13
hsa-mir-195	MIMAT0000461	29.53	26.16	5.29
hsa-mir-25	MIMAT0000081	28.92	25.53	5.29
hsa-mir-142-3p	MIMAT0000434	25.57	22.19	5.37
hsa-mir-34a	MIMAT0000255	29.09	25.73	5.38
hsa-mir-16	MIMAT0000069	25.57	22.17	5.41
hsa-mir-126	MIMAT0000445	29.71	26.27	5.57
hsa-mir-126*	MIMAT0000444	32.02	28.56	5.65
hsa-mir-26b*	MIMAT0004500	34.95	31.44	5.83
hsa-mir-548c-5p	MIMAT0004806	31.23	27.71	5.93
hsa-mir-15b	MIMAT0000417	28.21	24.59	6.29
hsa-mir-92a	MIMAT0000092	26.68	22.97	6.72
hsa-mir-150	MIMAT0000451	30.55	26.54	8.26
hsa-mir-768-3p	MI0005117	29.64	25.51	9.05
hsa-mir-223	MIMAT0000280	21.21	17.03	9.32
hsa-mir-628-5p	MIMAT0004809	34.44	30.14	10.09
hsa-mir-342-3p	MIMAT0000753	32.39	28.08	10.14
hsa-mir-193b	MIMAT0002819	31.13	26.67	11.23
hsa-mir-135b	MIMAT0000758	31.58	26.87	13.41
hsa-mir-532	MIMAT0002888	30.94	26.06	15.06
hsa-mir-144*	MIMAT0004600	32.92	26.76	36.57

* Pre-amplification of the miRNA cDNA was performed to improved the sensitivity of the qPCR array and increase the number of detectable miRNAs

Fig. 3 Regression plot analysis of miR-144* and miR-532-3p for biomarker validation, and miR-378 as a normalizer in feces of healthy and CRC subjects. **a** A nonsignificant positive correlation of miR-144* with miR-378 in healthy and CRC subjects was observed (CRC, $R^2 = 0.1835$ vs. healthy subjects, $R^2 = 0.3115$). **b** A significant positive correlation of miR-532-3p with miR-378 expression was found for the CRC group; R^2 was 0.59. **c** A significant correlation was not found between the expressions of miR-144* and miR-532-3p when analyzed together with miR-378. **d** Analysis of tissue specimens gave results similar to those obtained for miR-144* and miR-378 from fecal samples

Nevertheless, this confirmed that the expression of miR-144* in fecal miRNA is in accord with its expression in tissue samples (Fig. 3a, d). Furthermore, this data strengthens our conclusion that miRNA represents a novel, noninvasive, fecal-based diagnostic marker for CRC screening.
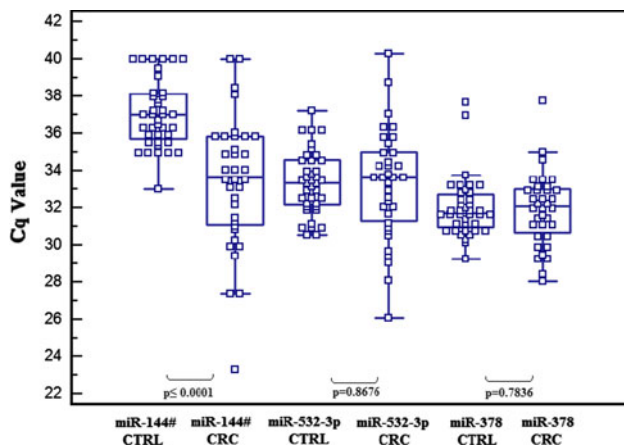
Fig. 4 Box and whisker plot analysis of miR-144*, miR-532-3p and miR-378 expression in feces of healthy and CRC patients. The direct C_q distribution of the normalizer miR-378 demonstrates that there was no significant difference between the expressions of miR-378 in diseased and healthy subjects (CRC $\bar{x}$ C_q expression: 31.92 ± 1.94 with; healthy subject $\bar{x}$ C_q expression: 32.02 ± 1.60), while the C_q distribution of miR-144* clearly distinguishes between CRC and healthy subjects with $p < 0.0001$. However, the direct C_q distribution of miR-532-3p indicates that there is no significant difference between the expression levels of miR-532-3p in healthy and CRC subjects (CRC $\bar{x}$ C_q expression: 33.27 ± 2.96 with; healthy subject $\bar{x}$ C_q expression: 33.34 ± 1.69), $p = 0.8676$

Discussion

Although miRNAs can be detected in many body fluids and in various tissues [15–23], only three reports have been published so far on the presence of miRNAs in feces [32–34]. Thus, further investigations of the properties of miR-NAs in feces are of the utmost importance. As we carried out our investigation, Ahmed et al. [32] demonstrated that the detection of specific miRNAs in fecal specimens clearly discriminated between sporadic human colon cancer and active ulcerative colitis in their patient cohort. Furthermore, Link et al. proposed the use of fecal miRNAs as novel biomarkers for CRC screening. They showed that miR-21 and miR-106a are highly overexpressed in adenomas and CRCs compared with individuals who do not have colorectal neoplasia [34]. On the other hand, Koga et al. [33] showed that the expressions of the miR-17-92 cluster and miR-135 were significantly higher in feces of CRC than healthy volunteers. Therefore, miRNA evaluation in feces is a promising and noninvasive approach to CRC screening, although further validation of pre-analytical sampling and miRNA stability are required. Thus, our analysis of the presence and stability of miRNAs and the identification of a putative candidate miRNA strengthens their conclusions. In this study, we mainly focused on determining miRNA stability in feces through methods such as the consistent detection of miRNAs using two different temperatures over time.

We have previously shown that long DNA arising from malignant colonocytes represents a fecal-based diagnostic marker of CRC [26, 27]. To further investigate the diagnostic power of fecal specimens as a noninvasive tool for CRC screening, we profiled feces for 648 miRNAs in a subset of CRC and healthy subjects. Based on our microarray data, we detected 39% of the miRNAs in the feces. To further analyze the stability of the miRNAs, we measured the expressions of endogenous miRNAs and exogenous cel-miR-238 in feces. Chim et al. [15] demonstrated that cell-free maternal and placental origin miRNAs were much more stable than exogenous miRNAs. Furthermore, Park et al. [17] reported that miRNAs could be detected in saliva samples and are much more stable than exogenous miRNAs, and a very slow degradation pattern was observed over time. Nevertheless, Tsui et al. [22] reported that bodily fluid miRNAs are more stable in their microenvironment. Taken together, our data confirm that miR-NAs are also stable in the fecal microenvironment.

Next, to evaluate whether miRNAs are stable at different temperatures, we accounted for the fact that fecal mRNA and DNA can be degraded very quickly due to the activities of bacterial RNase and DNase [35]. The complex nature of feces inhibits certain assays like mRNA quantification and/or protein detection for the purposes of disease discrimination. The transportation and processing of feces is a drawback, because the samples need to be stored in fecal stabilization buffer and/or handled as quickly as possible. Furthermore, this degradation is common regardless of the processing method employed, as shown by the reduced sizes of high molecular weight ribosomal bands and the presence of low molecular weight RNA. We found that samples kept for up to 1 year at $-20°C$ could be processed and quantified for miRNA detection (unpublished observation in our laboratory). Furthermore, we found that fecal miRNAs were much more stable at $4°C$ over a period of six days. Besides, Park et al. [17] showed that salivary miRNA could identify a diseased state in oral cancer detection. They reported that mature miRNAs are bound to a RISC complex, which showed that Ago 2 protein is also present in salivary specimens. This led us to hypothesize that miRNAs could also be associated with macromolecules present in feces, making them approximately as stable in feces as they are in other bodily fluids.

Based on this remarkable finding on the stability of miRNA in feces, we conducted RT-qPCR assays for a selected set of miRNA candidates for CRC diagnosis. Many miRNAs, such as miR-16 and miR-135b, showed upregulation, in accordance with the literature [36], but insignificant results were obtained when calculating the actual sensitivities of these markers in our preliminary investigation. We therefore selected miR-144* and miR-532-3p as markers for further evaluation. We also chose

Table 2 Demographics of the CRC and healthy individuals analyzed for hsa-miR-144* and hsa-378 expression in fecal specimens

Sex	Age	TNM stage	CRC site	¥Raw C_q of miR-144*	¥Raw C_q of miR-378
M	61	II	Sigmoid colon	33.11	30.55
F	68	NA	Rectal cancer	35.02	31.15
F	59	I	Rectal cancer	33.52	33.18
M	72	II	Rectal cancer	32.18	29.37
F	72	NA	Sigmoid cancer	38.45	31.95
F	44	III	Sigmoid cancer	36.04	32.24
M	75	II	Cecum cancer	31.02	31.42
F	74	0	Cecum cancer	34.05	31.46
M	58	II	Cecum cancer	35	37.73
F	85	NA	Rectal cancer	32.51	32.96
M	66	III	Sigmoid cancer	29.9	32.45
M	49	III	Rectal cancer	34.05	33.52
M	60	*	Rectal cancer	32.15	32
F	80	0	Rectal cancer	40	33.01
F	75	NA	Sigmoid cancer	35.95	33.51
F	78	Unclassified	Sigmoid colon	35.91	31.6
F	63	I	Sigmoid colon	35.84	31.09
F	82	0	Rectal cancer	29.42	29.24
F	71	NA	Rectal cancer	35.93	29.45
F	88	NA	Cecum cancer	23.3	29.9
M	65	NA	Sigmoid colon	30.83	29.89
F	55	II	Sigmoid colon	27.5	28.05
M	75	*	Rectal cancer	40	33.13
F	62	NA	Rectal cancer	33.08	32.55
M	79	II	Sigmoid cancer	27.36	28.42
M	67	0	Cecum cancer	29.9	32.96
F	85	NA	Rectal cancer	38.09	31.15
F	82	0	Cecum cancer	35.12	33.55
F	75	NA	Rectal cancer	34.87	32.51
M	60	NA	Right colon	38.15	34.6
F	68	NA	Sigmoid colon	33.64	35
M	81	III	Rectal cancer	31.27	32.08
M	84	III	Right colon	31.45	32.2
F	76	III	Left colon	30.25	30.45
M	58	NA	Right colon	35.85	32.93
CRC $n = 35$				$\bar{x}$ raw $C_q = 33.45$	$\bar{x}$ raw $C_q = 31.92$
15 males, median age 72 years, range 44–88 years				SD = 3.66	SD = 1.94
Healthy individuals $n = 40$				$\bar{x}$ raw $C_q = 36.99$	$\bar{x}$ raw $C_q = 32.02$
13 males, median age 61 years, range 32–79 years				SD = 1.74	SD = 1.60
Difference in C_q expression between CRC and healthy individuals				$\bar{x}$ raw C_q CRC-CTRL 3.54	$\bar{x}$ raw C_q CRC-CTRL 0.11
				*$p = 0.0001$	$p = 0.81$

Total patients analyzed $n = 75$, CRC $n = 35$, healthy subjects $n = 40$

NA and *: histology report not available, as the patient underwent chemoradiotherapy or surgery at a different location

¥ Geometric mean of C_q with $n = 2$

miR-378 as our quality control and as our reference miRNA for data normalization, because it showed a very low standard deviation in C_q values in our pooled analysis. We found that miR-378 is consistently expressed across healthy and neoplastic individuals. Moreover, we observed that miR-144* was significantly overexpressed in CRC

compared to normal controls. When we calculated the overall fold change between these two groups, we found that this miRNA is upregulated at least 11-fold, and is easily detected in fecal samples. To further strengthen
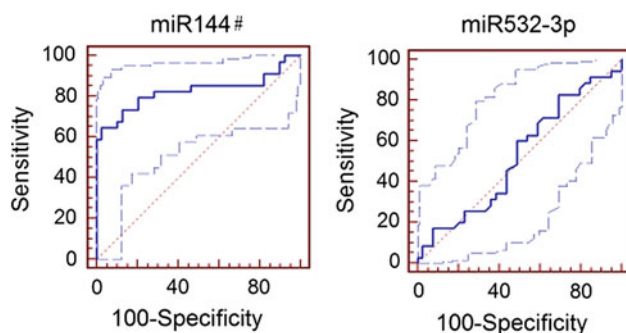

Fig. 5 Receiver operating characteristic (ROC) curve analyses of miR-144* and miR-532-3p in the fecal samples of a subset of CRC and healthy subjects, as derived from ΔC_q calculation. The *left-hand figure* shows the ROC analysis along with the corresponding area under the curve (AUC) for miR-144* expression. The ROC demonstrated a sensitivity of 74% and a specificity of 87%, AUC = 0.829. The *right-hand figure* shows the sensitivity and specificity of the miR-532-3p expression level when used to detect disease. *Solid line* represents the actual sensitivity and specificity of the test, while the *dashed lines* represent the 95% confidence interval

these data, the expression pattern of miR-144* was also analyzed in paired surgical tissues. As expected, we found that the expression was fivefold higher in CRC tissues compared to their surrounding tissues. These data further demonstrate that the levels of miRNA detected in feces strongly reflect the changes occurring in corresponding tissues. However, we observed a significant difference in the fold changes in the fecal and tumor samples (11- vs. 5-fold). This can be explained by the expression analysis, as it was done on two different subpopulations in our pilot study.

Finally, when we calculated the sensitivity and specificity of this assay, we demonstrated that miR-144* represents a novel fecal-based screening marker for CRC. Our findings are in agreement with the results reported by Koga et al. [33], with a sensitivity and a specificity of 74.1 and 79.0%, respectively.

Recently, Ahmed et al. [32] and Link et al. [34] showed that the expressions of miR-21 and miR-106a are strongly upregulated in feces of CRC patients. Their studies were also supported by Schetter and colleagues [14], who observed increased expressions of miR-21 and miR-106a in CRC and adenoma associated tissue from patients with poor survival and therapeutic outcomes. Our data from

Table 3 Demographics of CRC surgical tissues and the expression of hsa-miR-144*

Patients analyzed ($n = 15$ paired tissue samples)				Raw C_q of miR-144*		Raw C_q of miR-378		Fold change
Sex	Age	TNM stage	CRC sites	CRC	ST	CRC	ST	
F	46	NA	Undifferentiated carcinoma of the cecum	34.35	36.41	30.69	30.37	5.21
F	70	II	Left colon	35.17	36.31	31.03	30.74	2.69
F	76	III	Left colon	25.88	30.76	26.89	27.98	13.83
M	83	III	Right colon	33.34	28.91	30.61	29.34	0.11
M	74	II	Right colon	36.14	32.18	29.31	29.61	0.05
F	89	I	Cecum	31.00	33.56	29.35	28.42	11.24
F	62	III	Rectal	31.93	34.54	29.87	29.69	6.92
M	72	NA	Rectal	28.87	35.92	28.52	28.77	111.43
M	82	I	Rectal	28.86	30.54	28.64	26.86	11.00
M	79	III	Left colon	30.04	30.79	28.75	29.55	0.97
M	56	II	Rectal	29.50	33.60	30.50	29.95	25.11
M	70	III	Left colon	28.26	33.33	31.48	28.89	202.25
F	53	III	Rectal	29.59	27.52	30.87	27.74	2.08
M	58	I	Right colon	30.99	31.43	29.85	29.29	2.00
M	84	III	Right colon	29.16	32.99	28.58	28.79	12.30
9 Males, median age 72 years, range 46–89 years				$\bar{x}$ raw C_q = 30.87		$\bar{x}$ raw C_q = 29.66		
				SD = 2.83		SD = 1.24		
Paired surrounding tissues				$\bar{x}$ raw C_q = 32.59		$\bar{x}$ raw C_q = 29.07		
				SD = 2.64		SD = 1.03		
				Overall fold change		$2^{-\Delta\Delta C_q}$		4.99

ST surrounding tissues, *NA* not available

pooled analysis also confirmed these miRNA increases in CRC, with a borderline fold change for miR-21 (1.8-fold) and a significant fold change in the expression of miR-106a (threefold). Moreover, Xie et al. [23] demonstrated that the expression of miR-21 is significantly higher in sputum samples from patients with non-small cell lung cancer than in those from cancer-free individuals.

Moreover, our chosen miRNA, miR-144*, has not been reported to be correlated with CRC incidence, but we found significant upregulation of miR-144* in both feces and tissue samples from CRC patients. We found that miR-144* may be involved in programmed cell death mechanisms, as this miRNA targets APAF1, a cytoplasmic protein that initiates apoptosis (http://www.microrna.org/), and other potential targets and pathways that are yet to be identified.

The results of our study suggest that further experimentation in this area is needed. The use of a pooled sample approach for candidate miRNA screening is economically very attractive, but our data could be biased due to individual sample variation. Thus, in order to validate these findings, additional experiments using another set of patients who are undergoing CRC screening should be studied to determine the robustness of this approach. Moreover, the overexpression of miR-144* could reflect its involvement in the tumorigenesis of CRC. Therefore, further experimental analysis is needed to rule this out and establish the actual function of miR-144*, in order to show that miR-144* is a viable fecal-based diagnostic marker. The selection of an miRNA as an internal control in our study could be biased due to the unknown function of the miRNA itself. This selection could also differ from laboratory to laboratory and from study to study. Furthermore, in our own experience, endogenous control small RNAs (i.e., RNU19, U6) are not consistently expressed in cells, tissues, or even in fecal samples. Therefore, a better choice for a universal internal control for miRNA is necessary. Finally, an early diagnostic marker should emphasize the presence of premalignant CRC lesions. Thus, the detection of miR-144* in adenoma samples would be ideal, and would add further weight to the idea of using miRNAs as a screening modality.

This proof of principle study has revealed the potential of fecal miRNAs as practical molecular markers for diverse physiological and pathological conditions, such as cancer. We have demonstrated the ease and reliability with which fecal miRNAs can be determined, which may pave the way to their wider application as a possible noninvasive tool for CRC detection. Thus, large randomized control studies are urgently required to validate the results found here, and to discover other specific miRNAs that could represent the new generation of biomarkers for rapid diagnosis.

References

1. Cummins JM, He Y, Leary RJ, Pagliarini R, Diaz LA Jr, Sjoblom T, et al. The colorectal microRNAome. Proc Natl Acad Sci USA. 2006;103:3687–92.
2. Akao Y, Nakagawa Y, Naoe T. MicroRNA-143 and -145 in colon cancer. DNA Cell Biol. 2007;26:311–20.
3. Ventura A, Jacks T. MicroRNAs and cancer: short RNAs go a long way. Cell. 2009;136:586–91.
4. Aqeilan RI, Calin GA, Croce CM. miR-15a and miR-16-1 in cancer: discovery, function and future perspectives. Cell Death Differ. 2010;17:215–20.
5. Gartel AL, Kandel ES. miRNAs: little known mediators of oncogenesis. Semin Cancer Biol. 2008;18:103–10.
6. Fujita S, Iba H. Putative promoter regions of miRNA genes involved in evolutionarily conserved regulatory systems among vertebrates. Bioinformatics. 2008;24:303–8.
7. He L, Hannon GJ. MicroRNAs: small RNAs with a big role in gene regulation. Nat Rev Genet. 2004;5:522–31.
8. Koturbash I, Zemp FJ, Pogribny I, Kovalchuk O. Small molecules with big effects: the role of the microRNAome in cancer and carcinogenesis. Mutat Res. 2010;722:94–105.
9. Kloosterman WP, Plasterk RH. The diverse functions of microRNAs in animal development and disease. Dev Cell. 2006;11:441–50.
10. Stefani G, Slack FJ. Small non-coding RNAs in animal development. Nat Rev Mol Cell Biol. 2008;9:219–30.
11. Lu J, Qian J, Chen F, Tang X, Li C, Cardoso WV. Differential expression of components of the microRNA machinery during mouse organogenesis. Biochem Biophys Res Commun. 2005;334:319–23.
12. Ng EK, Chong WW, Jin H, Lam EK, Shin VY, Yu J, et al. Differential expression of microRNAs in plasma of patients with colorectal cancer: a potential marker for colorectal cancer screening. Gut. 2009;58:1375–81.
13. Calin GA, Ferracin M, Cimmino A, Di Leva G, Shimizu M, Wojcik SE, et al. A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N Engl J Med. 2005;353:1793–801.
14. Schetter AJ, Leung SY, Sohn JJ, Zanetti KA, Bowman ED, Yanaihara N, et al. MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. JAMA. 2008;299:425–36.
15. Chim SS, Shing TK, Hung EC, Leung TY, Lau TK, Chiu RW, et al. Detection and characterization of placental microRNAs in maternal plasma. Clin Chem. 2008;54:482–90.
16. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL, et al. Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci USA. 2008;105:10513–8.
17. Park NJ, Zhou H, Elashoff D, Henson BS, Kastratovic DA, Abemayor E, et al. Salivary microRNA: discovery, characterization, and clinical utility for oral cancer detection. Clin Cancer Res. 2009;15:5473–7.
18. Li J, Smyth P, Flavin R, Cahill S, Denning K, Aherne S, et al. Comparison of miRNA expression patterns using total RNA

extracted from matched samples of formalin-fixed paraffin-embedded (FFPE) cells and snap frozen cells. BMC Biotechnol. 2007;7:36.

19. Nelson PT, Baldwin DA, Kloosterman WP, Kauppinen S, Plasterk RH, Mourelatos Z. RAKE and LNA-ISH reveal microRNA expression and localization in archival human brain. RNA. 2006;12:187–91.

20. Xi Y, Nakajima G, Gavin E, Morris CG, Kudo K, Hayashi K, et al. Systematic analysis of microRNA expression of RNA extracted from fresh frozen and formalin-fixed paraffin-embedded samples. RNA. 2007;13:1668–74.

21. El-Hefnawy T, Raja S, Kelly L, Bigbee WL, Kirkwood JM, Luketich JD, et al. Characterization of amplifiable, circulating RNA in plasma and its potential as a tool for cancer diagnostics. Clin Chem. 2004;50:564–73.

22. Tsui NB, Ng EK, Lo YM. Stability of endogenous and added RNA in blood specimens, serum, and plasma. Clin Chem. 2002;48:1647–53.

23. Xie Y, Todd NW, Liu Z, Zhan M, Fang H, Peng H, et al. Altered miRNA expression in sputum for diagnosis of non-small cell lung cancer. Lung Cancer. 2009;67:170–6.

24. Ahlquist DA, Skoletsky JE, Boynton KA, Harrington JJ, Mahoney DW, Pierceall WE, et al. Colorectal cancer screening by detection of altered human DNA in stool: feasibility of a multitarget assay panel. Gastroenterology. 2000;119:1219–27.

25. Dong SM, Traverso G, Johnson C, Geng L, Favis R, Boynton K, et al. Detecting colorectal cancer in stool with the use of multiple genetic targets. J Natl Cancer Inst. 2001;93:858–65.

26. Kalimutho M, Del Vecchio Blanco G, Gravina P, Cretella M, Mannucci L, Mannisi E, et al. Quantitative denaturing high performance liquid chromatography (Q-dHPLC) detection of APC long DNA in faeces from patients with colorectal cancer. Clin Chem Lab Med. 2010;48:1303–11.

27. Kalimutho M, Del Vecchio Blanco G, Cretella M, Mannisi E, Sileri P, Formosa A, et al. A simplified, non-invasive fecal-based DNA integrity assay and iFOBT for colorectal cancer detection. Int J Colorectal Dis. 2011;26:583–92.

28. Taylor CF, Field D, Sansone SA, Aerts J, Apweiler R, Ashburner M, et al. Promoting coherent minimum reporting guidelines for biological and biomedical investigations: the MIBBI project. Nat Biotechnol. 2008;26:889–96.

29. Mestdagh P, Feys T, Bernard N, Guenther S, Chen C, Speleman F, et al. High-throughput stem-loop RT-qPCR miRNA expression profiling using minute amounts of input RNA. Nucleic Acids Res. 2008;36:e143.

30. Mestdagh P, Van Vlierberghe P, De Weer A, Muth D, Westermann F, Speleman F, et al. A novel and universal method for microRNA RT-qPCR data normalization. Genome Biol. 2009;10:R64.

31. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 2002;3:RESEARCH0034.

32. Ahmed FE, Jeffries CD, Vos PW, Flake G, Nuovo GJ, Sinar DR, et al. Diagnostic microRNA markers for screening sporadic human colon cancer and active ulcerative colitis in stool and tissue. Cancer Genomics Proteomics. 2009;6:281–95.

33. Koga Y, Yasunaga M, Takahashi A, Kuroda J, Moriya Y, Akasu T, et al. MicroRNA expression profiling of exfoliated colonocytes isolated from feces for colorectal cancer screening. Cancer Prev Res (Phila). 2010;3:1435–42.

34. Link A, Balaguer F, Shen Y, Nagasaka T, Lozano JJ, Boland CR, et al. Fecal MicroRNAs as novel biomarkers for colon cancer screening. Cancer Epidemiol Biomarkers Prev. 2010;19:1766–74.

35. Yu YJ, Majumdar AP, Nechvatal JM, Ram JL, Basson MD, Heilbrun LK, et al. Exfoliated cells in stool: a source for reverse transcription-PCR-based analysis of biomarkers of gastrointestinal cancer. Cancer Epidemiol Biomarkers Prev. 2008;17:455–8.

36. Bartels CL, Tsongalis GJ. MicroRNAs: novel biomarkers for human cancer. Clin Chem. 2009;55:623–31.