Identification of a Five-MiRNA Expression Assay to Aid Colorectal Cancer Diagnosis

Matthew G. Davey 1,2,*, Gerard Feeney 1, Heidi Annuk 1, Maxwell Paganga 3, Emma Holian 3, Aoife J. Lowery 1, Michael J. Kerin 1 and Nicola Miller 1

1 Department of Surgery, Lambe Institute for Translational Research, National University of Ireland, H91 TK33 Galway, Ireland
2 Department of Surgery, Galway University Hospitals, H91 YR71 Galway, Ireland
3 School of Mathematical and Statistical Sciences, National University of Ireland, H91 TK33 Galway, Ireland
* Correspondence: m.davey7@nuigalway.ie

Abstract: Introduction: One-third of colorectal cancer (CRC) patients present with advanced disease, and establishing control remains a challenge. Identifying novel biomarkers to facilitate earlier diagnosis is imperative in enhancing oncological outcomes. We aimed to create miRNA oncogenic signature to aid CRC diagnosis. Methods: Tumour and tumour-associated normal (TAN) were extracted from 74 patients during surgery for CRC. RNA was isolated and target miRNAs were quantified using real-time reverse transcriptase polymerase chain reaction. Regression analyses were performed in order to identify miRNA targets capable of differentiating CRC from TAN and compared with two endogenous controls (miR-16 and miR-345) in each sample. Areas under the curve (AUCs) in Receiver Operating Characteristic (ROC) analyses were determined. Results: MiR-21 (β-coefficient:3.661, SE:1.720, p = 0.033), miR-31 (β-coefficient:2.783, SE:0.918, p = 0.002), and miR-150 (β-coefficient:−4.404, SE:0.526, p = 0.004) expression profiles differentiated CRC from TAN. In multivariable analyses, increased miR-31 (β-coefficient:2.431, SE:0.715, p < 0.001) and reduced miR-150 (β-coefficient:−4.620, SE:1.319, p < 0.001) independently differentiated CRC from TAN. The highest AUC generated for miR-21, miR-31, and miR-150 in an oncogenic expression assay was 83.0% (95%CI: 61.7–100.0, p < 0.001). In the circulation of 34 independent CRC patients and 5 controls, the mean expression of miR-21 (p = 0.001), miR-31 (p = 0.001), and miR-150 (p < 0.001) independently differentiated CRC from TAN. The highest AUC generated for miR-21, miR-31, and miR-150 in an oncogenic expression assay was 83.0% (95%CI: 61.7–100.0, p < 0.001). In the circulation of 34 independent CRC patients and 5 controls, the mean expression of miR-21 (p = 0.001), miR-31 (p = 0.001), and miR-150 (p < 0.001) differentiated CRC from controls; however, the median expression of miR-21 (p = 0.476), miR-31 (p = 0.933), and miR-150 (p = 0.148) failed to differentiate these groups. Conclusion: This study identified a five-miRNA signature capable of distinguishing CRC from normal tissues with a high diagnostic test accuracy. Further experimentation with this signature is required to elucidate its diagnostic relevance in the circulation of CRC patients.

Keywords: colorectal cancer; miRNA; genomics; cancer diagnostics; personalised medicine

1. Introduction

In the Western world, colorectal cancer (CRC) is the second most common cause of cancer-related death [1]. Earlier CRC detection through screening programs, novel multimodal treatment strategies, and an increased understanding of the molecular and hereditary processes driving the disease have all contributed to improved clinical outcomes for patients diagnosed with CRC [2]. Despite these marked improvements, 30% of patients present with distant metastasis and will still succumb to their disease, with 5-year survival rates ranging from 90% in the setting of stage 1 cancers to approximately 10% in patients with stage IV cancers [3,4]. Unfortunately, many patients with CRC remain asymptomatic until their disease is advanced, rendering cure a typically unsurmountable challenge. At present, clinical practice is heavily dependent upon the carcinoembryonic antigen (CEA) biomarker, the clinical uses of which include post-operative surveillance, gauging
response to treatment, and estimating prognoses [5]. Despite this, reliance upon CEA as a diagnostic biomarker is limited by moderate sensitivity levels [6] (particularly in the setting of early-stage disease [7]), which suggests that CEA is unsuitable at the population level to be utilised as an adjunct to aid screening for CRC [8–11]. Thus, it is imperative that translational research efforts are focused upon identifying novel diagnostic biomarkers to facilitate earlier detection and intervention for patients diagnosed with CRC.

Microribonucleic acids (miRNAs) are small, non-coding ribonucleic acids (RNA) approximately 19–25 nucleotides in length that are crucial in the regulation of gene expression [12]. MiRNAs are estimated to regulate up to 30% of the human genome [13] through the alteration of genetic expression at a post-transcriptional level by acting on specific messenger RNA (mRNA) targets, inducing mRNA degradation or translational inhibition [14,15]. Aberrant miRNA expression profiles have been implicated as key regulators in cancer proliferation and metastasis in malignancies such as CRC [16–18], with potential prognostic, diagnostic, and therapeutic avenues being explored through miRNA evaluation [19–21]. For example, the measurement of miR-135b and miR-195 in patients diagnosed with CRC has recently been correlated with long-term oncological and survival outcomes [22]. Moreover, it is now well established that miRNAs maintain stability in an array of biological tissues (including tumour tissue, ‘normal’ epithelium, and circulation) and may be quantified relatively simply and inexpensively using real-time quantitative reverse transcriptase polymerase chain reaction (RT-qPCR) [19,23,24]. While the extrapolation of miRNAs which may inform CRC diagnosis in the clinical setting is imperative, the identification of such biomarkers is primarily reliant on those expressed differentially in tumour tissue and in patient circulation. Accordingly, the aim of the current study was to identify miRNAs capable of distinguishing colorectal tumour tissue from tumour-associated ‘normal’ (or TAN) control tissue and to create a miRNA oncogenic signature to aid in CRC diagnosis. We then aimed to determine the value of measuring these miRNAs in the circulation of an independent cohort of CRC patients and ‘normal’ controls to help inform CRC diagnosis. In this study, we investigated the expression of a panel of 17 miRNA targets in a cohort of tumour and normal tissue. We then determined which miRNA targets most significantly differentiated CRC tissue from TAN before including these in a five-miRNA signature, with two validated endogenous controls. We then evaluated the use of these miRNAs as diagnostic biomarkers in the circulation of independent patient samples.

2. Materials and Methods

Local hospital ethical approval was obtained from the Galway Clinical Ethical Research Committee from Galway University Hospitals (C.A. 45/05 and C.A. 151). Informed and written consent was obtained from all included patients before tumour and tumour-associated normal (TAN) tissue samples were obtained from a cohort of 74 consecutive patients being treated surgically with curative intent for CRC at Galway University Hospitals. TAN tissue was obtained from the ‘normal’ epithelium adjacent to the cancer at the time of resection for use as control tissue in the training analysis. Furthermore, liquid biopsy in the form of venous sampling was performed on an independent cohort of 34 CRC patients and 5 ‘normal’ controls. These were used as a validation set to determine the role of miRNA targets in distinguishing patients with CRC from ‘healthy’ controls. All patient demographic and clinicopathological data were obtained from a prospectively maintained database at the Department of Surgery at the National University of Ireland, Galway. This study was performed in accordance with the STROBE guidelines for cohort studies.

2.1. Tumour Staging

All the consecutive 74 patients included in this analysis had previously presented to our local tertiary referral centre at Galway University Hospitals for the multidisciplinary management of their CRC in accordance with standard-of-care guidelines. Galway University Hospitals is a hospital providing cancer services to a population of almost 900,000 living in the west of Ireland. Each patient had to have prior histopathological confirmation of
CRC at their local accredited histopathological laboratory (or satellite laboratories) with staging performed in accordance with tumour, node, and metastasis (TNM) staging profile in accordance with the American Joint Committee on Cancer (AJCC) version 8 [23]. Tumour staging was performed using a combination of clinical and radiological staging prior to resection, before histopathological evaluation provided definitive tumour staging based on resected specimens.

2.2. Immunohistochemical Tumour Evaluation

Conventional immunophenotypical staining using cytokeratin (CK) 20 positivity and CK7 negativity was used to discriminate adenocarcinoma from other histological colorectal subtypes [24]. Thereafter, CDX2 was used to determine tumour differentiation [25]. Thereafter, the presence of lymphatic invasion was evaluated using D2-40 staining and vascular invasion using CD34 (combined, these are typically reported as lymphovascular invasion or LVI) [26,27]. Simultaneous appraisal of tumour perineural invasion (PNI) was evaluated using S-100 staining [28] and extramural vascular invasion (EMVI) was obtained using elastin staining [29]. These were performed on all resected specimens in accordance with standard-of-care guidelines.

2.3. Radiological Staging

Prior to CRC resections, the radiological staging of each tumour was performed. This involved using computed tomography (CT) for all cases of CRC. The CT scanners used to perform this staging were Siemens Somatom Definition AS 128 Slice CT scanners. In cases of neoplasms of the rectum, additional pelvic magnetic resonance imaging (MRI) was used to further evaluate and stage the extent of disease in the soft tissue of the pelvis. The MRI scanner used was a short bore 1.5 T magnet (Magnetom Espree 1.5 T, Siemens Healthcare, Erlangen, Germany). Available clinical and pathological data for the 74 included patients are outlined in detail in Table 1.

Table 1. Clinicopathological data of the 74 patients with colorectal cancer included in this study.

<table>
<thead>
<tr>
<th>Clinicopathological Parameter</th>
<th>Patients with Colorectal Cancer (N = 74)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mean age</strong> (±standard deviation, range)</td>
<td>67.8 years (±12.5, 38–90 years)</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>- Male</td>
<td>51 (68.9%)</td>
</tr>
<tr>
<td>- Female</td>
<td>23 (31.1%)</td>
</tr>
<tr>
<td>Tumour Location</td>
<td></td>
</tr>
<tr>
<td>- Colon</td>
<td>52 (70.3%)</td>
</tr>
<tr>
<td>- Rectum</td>
<td>22 (29.7%)</td>
</tr>
<tr>
<td>Presentation</td>
<td></td>
</tr>
<tr>
<td>- Emergency</td>
<td>12 (16.2%)</td>
</tr>
<tr>
<td>- Elective</td>
<td>62 (83.8%)</td>
</tr>
<tr>
<td>Histological subtype</td>
<td></td>
</tr>
<tr>
<td>- Adenocarcinoma</td>
<td>53 (68.0%)</td>
</tr>
<tr>
<td>- Mucinous</td>
<td>3 (3.8%)</td>
</tr>
<tr>
<td>- Other/Missing</td>
<td>22 (28.2)</td>
</tr>
<tr>
<td>Tumour Stage</td>
<td></td>
</tr>
<tr>
<td>- T1</td>
<td>2 (2.7%)</td>
</tr>
<tr>
<td>- T2</td>
<td>5 (6.8%)</td>
</tr>
<tr>
<td>- T3</td>
<td>25 (33.8%)</td>
</tr>
<tr>
<td>- T4</td>
<td>18 (24.3%)</td>
</tr>
<tr>
<td>- TX</td>
<td>24 (32.4%)</td>
</tr>
<tr>
<td>Nodal Stage</td>
<td></td>
</tr>
<tr>
<td>- N0</td>
<td>15 (20.3%)</td>
</tr>
<tr>
<td>- N1</td>
<td>20 (27.0%)</td>
</tr>
<tr>
<td>- N2</td>
<td>10 (13.5%)</td>
</tr>
<tr>
<td>- NX</td>
<td>29 (39.2%)</td>
</tr>
</tbody>
</table>
2.4. Calculating Follow-Up

For each of the 74 included patients, follow-up was recorded through a prospectively maintained institutional database, which was stored at the Department of Surgery at the National University of Ireland, Galway. Median lengths of follow-up were calculated using the reverse Kaplan–Meier method, as described previously by Xue et al. [30]. All clinicopathological, surgical, and follow-up data were cross-referenced using patient electronic and hard copy medical records to ensure accuracy.

2.5. Identification of MiRNA Targets

A pre-determined miRNA panel of 15 cancer-associated miRNAs (miR-17, miR-20a, miR-21, miR-31, miR-132, miR-135b, miR-139-5p, miR-145, miR-148a, miR-150, miR-155, miR-195, miR-200c, miR-203, and miR-215) was identified during an extensive review of the previous reported literature. These miRNAs were selected for inclusion in our study based on their relevance to tumours of the colon, rectum, and other epithelial cancers (Table 2) [12,20,31–43]. In this analysis, the discovery cohort included approximately one-third of the entire sample cohort (27 paired tumour and TAN samples—36.5%), which were randomly selected. Following the results in the discovery cohort, the expression levels of 4 miRNAs (miR-31, miR-135b, miR-150, and miR-155) were successfully identified to differentiate tumour from TAN tissues and were then considered suitable for inclusion in the validation cohort of 47 paired CRC tissue and TAN tissue for analysis (63.5%). Two other miRNA targets (miR-21 and miR-195) were also chosen based on previous work from our group demonstrating the oncogenic properties of these miRNAs in cancer [12,20] and investigated in this validation analysis.

Table 2. The relevance of the use of discovery cohort of 15 target miRNA and 2 endogenous controls in differentiating tumour tissue from tumour-associated normal tissue in 27 colorectal cancer tissue and 27 ‘normal’ tissues.
Table 2. Cont.

<table>
<thead>
<tr>
<th>Target</th>
<th>MiRNA Function</th>
<th>Expression Levels</th>
<th>CT Difference</th>
<th>Efficiencies</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-195</td>
<td>Known oncogenic biomarker in malignancy (12)</td>
<td>Decreased</td>
<td>11.88</td>
<td>93%</td>
<td>0.245</td>
</tr>
<tr>
<td>miR-200c</td>
<td>Regulator of metastasis within CRC (40)</td>
<td>Decreased</td>
<td>11.88</td>
<td>101%</td>
<td>0.323</td>
</tr>
<tr>
<td>miR-203</td>
<td>Diagnostic and prognostic biomarker in CRC (41)</td>
<td>Increased</td>
<td>12.84</td>
<td>104%</td>
<td>0.146</td>
</tr>
<tr>
<td>miR-215</td>
<td>Prognostic biomarker in CRC (42)</td>
<td>Decreased</td>
<td>10.46</td>
<td>93%</td>
<td>0.139</td>
</tr>
<tr>
<td>miR-16</td>
<td>Endogenous control (44)</td>
<td>Stable</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>miR-345</td>
<td>Endogenous control (44)</td>
<td>Stable</td>
<td>0.00</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

CT, cycle threshold; CRC, colorectal cancer. (□) denotes independent samples t-test (□). * denotes statistical significance.

2.6. RNA Isolation and Biobanking

The extraction of RNA was performed using MagNA Pure Isolation (Roche), as per the manufacturers’ recommendations. In brief, total RNA extraction was performed on 1µL of tumour/TAN homogenate tissue prior to the RNA concentrations and integrity being formally determined using NanoDrop® spectrophotometry (Nanodrop ND-1000 Technologies Inc., Wilmington, DE, USA) and the Agilent Bioanalyser (Agilent Technologies, Germany). For the extraction of RNA from liquid biopsies (blood), an amount of 1 µL was also utilised. RNA concentrations were thus determined and their associated 260/230 and 260/280 ratios were recorded (with values within the target range of 2.0–2.2 being deemed acceptable). Thereafter, the RNA integrity was evaluated using RNA 6000 Nano LabChip Series II Assays (for small RNA) performed on a 2100 Bioanalyzer (Agilent Technologies, Waldbronn, Germany). The yielded RNA was then transferred to storage tubes and labelled in a pseudoanonymised fashion. These samples were then stored at −70 °C in the local Cancer Biobank at the Department of Surgery at the National University of Ireland, Galway.

2.7. Efficiency Calculations

The percentage of polymerase chain reaction (PCR) amplification efficiencies (E) for each miRNA target was calculated using the slope of the semi-log regression plot of cycle threshold vs. log input of cDNA (10-fold dilution series of five points) with the following equation, and a threshold of 10% above or below 100% efficiency was applied. For efficiency calculations, the miRNA targets identified in the literature review were used to ensure the scientists’ competency with all targets of interest prior to commencing miRNA analysis on CRC and TAN tissue. The efficiency equation used was as follows:

\[ E = (10^{-1/slope} - 1) \times 100 \]  

Once efficiency was achieved for each target, miRNA analysis was performed using RT-qPCR.

2.8. Reverse Transcription Polymerase Chain Reactions

Thereafter, stored RNA samples were retrieved from the biobank repository and subjected to reverse transcription using miRNA primers. These miRNA primers were TaqMan® assays specific to each miRNA (Applied Biosystems, Foster City, CA, USA), and RT-qPCR was then performed as per the manufacturer’s instructions.

In brief, this involved performing the reverse transcription of 5 ng of tumour/TAN (or 100 ng of blood) total RNA using the MultiScribe™ based High-Capacity cDNA Archive kit (Applied Biosystems, Foster City, CA, USA). In these analyses, reverse transcriptase controls were included to ensure consistent normalisation across each reaction. Thereafter,
PCR tests were performed in final volumes of 10 µL using the QuantStudio 7 Flex Fast Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). These reactions were performed on 1.0 µL cDNA, 5 µL TaqMan® Universal PCR Fast Master Mix, 3.5 µL of nuclease-free water, and 0.5 µL TaqMan® primer–probe mix (Applied Biosystems, Foster City, CA, USA). These reactions were initiated with a 10 min incubation period at 95 °C followed by 40 cycles at 95 °C for 15 s and 60 °C for 60 s.

2.9. Endogenous and Inter-Assay Controls

MiRNA target miR-26b was used as an inter-assay control from tissue derived from the MDA-MB-468 breast cancer cell line. This control was included on each plate to ensure normalisation across plates, and all reactions were performed in triplicate to ensure the consistency and robustness of the results. Based on previously reported findings, miR-16 and miR-345 were used as endogenous controls to ensure the normalisation and standardisation of miRNA expression within CRC and TAN tissues [44,45]. The threshold standard deviation (SD) for intra-assay and inter-assay replicates was set as 0.3. Following this, the miRNA expression levels were calculated using the QbasePlus software (Biogazelle, Gent, Belgium). The geNorm method was used to ensure that the results were calibrated and normalised. The MiRNA expression profiles were then relatively quantified compared to the pre-determined endogenous controls (miR-16 and miR-345) to ensure normalisation and standardisation across samples [45,46].

2.10. Statistical Evaluation

For each miRNA target, the expression levels within tumour and TAN samples were expressed as means (with their associated standard error (SE)), and the means were compared using a paired-samples t-test (□). The comparison of medians was then performed using the Wilcoxon signed-rank test (†). Univariable and multivariable logistic and linear regression analyses were performed on miRNA expression profiles to determine the ability of such miRNAs to differentiate tumour tissues from TAN. The results were expressed as β-coefficients with associated SEs and 95% confidence intervals (CIs). Values with $p < 0.050$ following univariable analysis were considered for inclusion in the multivariable analysis. Following this, regression trees were used to classify the clinically relevant cut-offs for each miRNA included in the multivariable regression analysis. Receiver Operating Characteristic (ROC) curve analyses were carried out using binary logistic regression analysis, with area under the curve (AUC), sensitivity, and specificity used to express the diagnostic test accuracy. Our cohort was randomly divided into a test set (79.7%, 59/74) and a validation set (20.3%, 15/74). All tests of significance were 2-tailed, with $p < 0.050$ indicating statistical significance. Differential miRNA expression was expressed as a fold change calculated by log$_2$ (RQ$_{\text{Tumour}}$/RQ$_{\text{TAN}}$). Statistical analysis was performed using the statistical package for social sciences (SPSS) version 26.0 (Chicago, IL, USA), and data were analysed using the R statistical software version 3.2.3 (Auckland, New Zealand).

3. Results

3.1. Included Patients

A total of 74 patients diagnosed with colorectal cancer donated tumour tissue for use in this study. Clinicopathological data for these patients are shown in Table 1. TAN tissue was available from 74 of these included patients. The median follow-up period was 85.6 months ($\pm$ 9.5 months).

3.2. Target miRNA as Oncogenic Biomarkers in Colorectal Cancer

In the discovery cohort, we investigated the expression of 15 miRNAs (miR-17, miR-20a, miR-21, miR-31, miR-132, miR-135b, miR-139-5p, miR-143, miR-145, miR-148a, miR-150, miR-155, miR-195, miR-200c, miR-203, and miR-215) (Table 2). Of these, four miRNAs (miR-31 miR-135b, miR-150, and miR-155) were found to be able to significantly differentiate tumour tissue from TAN (all $p < 0.050$, ¶). In the validation cohort, two other miRNAs
were chosen for investigation based on previous work from our group demonstrating the use of miR-21 and miR-195 as oncogenic miRNAs within the setting of cancer [12,20].

The aberrant expression of miR-21, miR-31, miR-135b, and miR-150 was found to be able to significantly differentiate tumour tissue from TAN (all \( p < 0.050 \)), corresponding to an increase in fold change for miR-21, miR-31, and miR-135b and a decrease in fold change for miR-150 (Table 3 and Figure 1). MiRNA expression profiles found to be associated with differentiating tumour and TAN were then included in linear regression analyses. Using univariable logistic regression analyses, the expression profiles of miR-21, miR-31, and miR-150 were found to be able to differentiate tumour tissue from TAN (all \( p < 0.050 \)). In multivariable analyses, the expression profiles of miR-31 (\( \beta \)-coefficient: 2.431; SE: 0.715, \( p < 0.001 \)) and miR-150 (\( \beta \)-coefficient: −4.620; SE: 1.319, \( p < 0.001 \)) were found to be able to independently differentiate colorectal tumour from TAN (Table 4).

![Figure 1](image-url)  
**Figure 1.** Boxplot analyses illustrating the difference in the mean fold change of miRNA expression in the combined data from discovery and validation sets for comparing tumour samples and tumour-associated normal tissue samples for the six relevant miRNAs.
Table 3. Fold change expression of miRNA targets in the combined discovery and validation sets used for discriminating colorectal tumour tissue from tumour-associated normal samples.

<table>
<thead>
<tr>
<th>Target</th>
<th>Tumour Mean Log₂ Fold Change</th>
<th>TAN Mean Log₂ Fold Change</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-21</td>
<td>0.23 (SE: 0.47)</td>
<td>-0.43 (SE: 0.53)</td>
<td>&lt;0.001 *†</td>
</tr>
<tr>
<td>miR-31</td>
<td>0.48 (SE: 0.73)</td>
<td>-0.54 (SE: 0.49)</td>
<td>&lt;0.001 *‡</td>
</tr>
<tr>
<td>miR-135b</td>
<td>0.46 (SE: 0.55)</td>
<td>-0.26 (SE: 0.63)</td>
<td>&lt;0.001 *†</td>
</tr>
<tr>
<td>miR-150</td>
<td>-0.17 (SE: 0.46)</td>
<td>0.22 (SE: 0.27)</td>
<td>&lt;0.001 *†</td>
</tr>
<tr>
<td>miR-155</td>
<td>0.04 (SE: 0.37)</td>
<td>0.01 (SE: 0.27)</td>
<td>0.312 †</td>
</tr>
<tr>
<td>miR-195</td>
<td>0.17 (SE: 0.60)</td>
<td>0.02 (SE: 0.83)</td>
<td>0.090 †</td>
</tr>
</tbody>
</table>

TAN, tumour-associated normal tissue; SE, standard error. † denotes Wilcoxon signed-rank test. * denotes statistical significance.

Table 4. Logistic regression univariable and multivariable analyses were used to determine the predictive role of miRNA targets in differentiating tumour samples from tumour-associated normal samples.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>β-Coefficient (SE) Univariable</th>
<th>p-Value</th>
<th>β-Coefficient (SE) Multivariable</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>MiR-21</td>
<td>3.661 (1.720)</td>
<td>0.033 *</td>
<td>2.431 (0.715)</td>
<td>&lt;0.001 *</td>
</tr>
<tr>
<td>MiR-31</td>
<td>2.783 (0.918)</td>
<td>0.002 *</td>
<td>2.431 (0.715)</td>
<td>&lt;0.001 *</td>
</tr>
<tr>
<td>MiR-135b</td>
<td>0.155 (0.882)</td>
<td>0.861</td>
<td>2.431 (0.715)</td>
<td>&lt;0.001 *</td>
</tr>
<tr>
<td>MiR-150</td>
<td>-4.404 (0.526)</td>
<td>0.004 *</td>
<td>-4.620 (1.319)</td>
<td>&lt;0.001 *</td>
</tr>
<tr>
<td>MiR-155</td>
<td>2.850 (2.960)</td>
<td>0.336</td>
<td>2.431 (0.715)</td>
<td>&lt;0.001 *</td>
</tr>
<tr>
<td>MiR-195</td>
<td>-0.694 (1.017)</td>
<td>0.495</td>
<td>2.431 (0.715)</td>
<td>&lt;0.001 *</td>
</tr>
</tbody>
</table>

SE, standard error; 95% CI, 95% confidence interval. * denotes statistical significance.

3.3. Development and Validation of a Five-miRNA Oncogenic Signature

Based on the results of our multivariable analysis, a logistic regression tree analysis was performed to include the two-target miRNA of interest (miR-31 and miR-150). This analysis classified the relevant clinical cut-offs for these miRNAs in the differentiation of CRC from TAN (Figure 2). Thereafter, AUC generated from the ROC curve analysis was generated using binary logistic regression analysis to include miR-21, miR-31, and miR-150. The highest AUC generated was 83.0% (95% CI: 61.7–100.0, p < 0.001) for miR-21, miR-31, and miR-150 in the validation set (20.3%, 15/74) (Figure 3).

We then attempted to determine the diagnostic capabilities of miR-21, miR-31, and miR-150 in the circulation of patients with colorectal cancer (N = 34) compared to ‘normal’ controls (N = 5). In the circulation, the mean expression of miR-21 (p = 0.001), miR-31 (p = 0.001), and miR-150 (p < 0.001) was able to successfully differentiate the independent cohort of 34 colorectal cancer patients from the 5 ‘normal’ controls. However, the median expression levels of miR-21 (p = 0.476), miR-31 (p = 0.933), and miR-150 (p = 0.148) failed to accurately differentiate these groups (Table 5). The ROC analyses used for independent miRNAs are outlined in the Supplementary Material.

Table 5. Comparison of the mean and median expression levels of miRNA targets in the circulation of 34 independent colorectal cancer patients and 5 ‘normal’ controls.

<table>
<thead>
<tr>
<th>Target</th>
<th>Cancer Patient Mean Log₂ Fold Change</th>
<th>p-Value</th>
<th>Cancer Patient Median Log₂ Fold Change</th>
<th>Control Mean Log₂ Fold Change</th>
<th>p-Value</th>
<th>Control Median Log₂ Fold Change</th>
<th>p-Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>miR-21</td>
<td>-0.64</td>
<td>0.001 *</td>
<td>0.25</td>
<td>0.42</td>
<td>0.001 *</td>
<td>1.02</td>
<td>0.476 †</td>
</tr>
<tr>
<td>miR-31</td>
<td>0.01</td>
<td>0.001 *</td>
<td>0.23</td>
<td>-0.05</td>
<td>0.001 *</td>
<td>0.25</td>
<td>0.933 †</td>
</tr>
<tr>
<td>miR-150</td>
<td>-0.10</td>
<td>0.001 *</td>
<td>0.18</td>
<td>0.69</td>
<td>&lt;0.001 *</td>
<td>1.03</td>
<td>0.148 †</td>
</tr>
</tbody>
</table>

* denotes statistical significance. † denotes independent samples t-test. ‡ denotes Wilcoxon signed-rank test.
Figure 2. Regression tree illustrating the clinical use of miR-31 and miR-150 measurement to differentiate tumour tissue from tumour-associated normal tissue.

Figure 3. Use of Receiver Operating Characteristic curve analysis for miR-21, miR-31, and miR-150 in differentiating tumour tissues from tumour-associated normal tissues in our validation set.

4. Discussion

The clinical management paradigm of CRC continues to weigh heavily upon CEA as a biomarker to ascertain early detection for those with undiagnosed CRC recurrence, despite the several well-reported shortcomings of the use of this biomarker in reliably detecting CRC [6,7,47]. In the current analysis, we successfully demonstrated the clinical utility of a novel five-miRNA signature capable of differentiating colorectal tumour tissue from ‘normal’ tissue, with strong diagnostic accuracy (AUC: 83.0% (95% CI: 61.7–100.0%) for miR-21, miR-31, and miR-150). This is particularly promising when compared with data from a recent systematic review by Shinkins et al. which estimated CEA sensitivity levels
to be approximately 50% in making correct CRC diagnoses [47]. Moreover, we successfully expressed these biomarkers in the circulation within an independent validation cohort; this shows promise in the translational research efforts to enhance current diagnostics for these patients to expedite CRC diagnoses, surgical resection, and multimodal treatment [48]. While these findings are encouraging, the authors acknowledge that this study is ultimately limited by the median measurements of these biomarkers in circulation failing to differentiate CRC patients from ‘normal’ controls in the validation cohort, which speaks to the clinical challenge posed by miRNA quantification in circulation [49]. Thus, the further interrogation of the scientific method is required to refine the discovery of novel diagnostic biomarkers in order to enhance oncological outcomes for our prospective patients facing CRC diagnoses.

In the current study, the aberrant expression of three miRNAs (miR-21, miR-31, and miR-150) was able to independently differentiate CRC tumour tissue from TAN specimens and the blood of CRC patients from ‘healthy’ controls, supporting their inclusion in our five-miRNA signature with two endogenous controls. These results are novel in that these data have never before been compiled and used to aid CRC diagnosis; however, they may be perceived to be somewhat unsurprising. Several studies have outlined the importance of miR-21 expression as a reliable and informative biomarker in oncogenesis, with several studies correlating miR-21 with proliferation, prognosis, and survival in CRC [43,50–52]. Moreover, miR-21 has been long established as an ‘oncomir’ in several epithelial cancers, including breast, colorectal, and esophageal carcinoma [43,53,54], making it unsurprising that miR-21 was found to promote differentiation between CRC tumour and TAN in the current study. Additionally, Slaby et al. previously described the upregulation of miR-31 in CRC tissue versus ‘normal’ controls [43], while several other reports correlated miR-31 expression with aggressive clinicopathological features and advanced tumour stage in CRC [50,55,56]. The gene encoding miR-31 gene is located on chromosome 9p21.3, in close proximity to the locus of the well-described tumour suppressors cyclin-dependent kinase inhibitor (CDKN)2A and CDKN2B, known to encode for the cell cycle inhibitor proteins p15 and p16 [57]. Given their proximity to these loci, it is reasonable to suggest that miR-31 could be dysregulated together with CDKN2A in various cancers [58]. This provides a rationale for miR-31 being at the epicentre of CRC oncogenesis as the most important miRNA for delineating tumour from TAN in this study. This is exemplified by its position at node 1 in our regression tree, which highlights its crucial role in dictating CRC diagnosis in our analysis of 148 tissues (74 CRC and 74 TAN). This result, reported in tandem with a ROC diagnostic test accuracy of 85.9% (95% CI: 79.3–92.4%) for miR-31, illustrates the potential value of miR-31 measurement and assessment in future clinical translational research studies. Additionally, the reduced expression of miR-150 was found to be successful in delineating CRC tissue from TAN in this study, a finding similar to that obtained previously in the work of Aherne et al. [38]. Increased miR-150 expression has tumour suppressor function in human colorectal cancer through the inhibition of c-Myb, a 75–80 kd transcription factor [59,60], which leads to a stepwise increase in expression levels from non-cancerous ‘normal’ tissues to benign polyps and to CRC tissue [61]. This makes it unsurprising that reduced miR-150 expression has been previously correlated with poorer clinical and survival outcomes in CRC [62–64] and that its reduced expression was able to discriminate CRC tissue from controls in our multivariable analysis, supporting its inclusion in our five-miRNA signature. Moreover, the inclusion of miR-150 as the sole associate of miR-31 in our two-miRNA signature exemplifies its relevance as an oncogenic biomarker which may inform diagnoses if applied appropriately. Similar to the promising results observed for miR-31, the ROC diagnostic test accuracy of miR-150 was 80.8% (72.8–88.9%), indicating its potential as a diagnostic oncological biomarker in the setting of CRC. Accordingly, this study illustrates the potential value of integrating miR-21, miR-31, and miR-150 into a novel five-miRNA expression assay which may support the diagnosis of CRC for prospective patients.
The current study adds to the body of evidence supporting the novel discovery and clinical utility of miRNAs as biomarkers capable of informing diagnoses within the setting of cancer [12,19,55,56,65–69]. However, despite these promising results, we must acknowledge that there has only been marginal advancements in our understanding of the biomolecular pathways which incorporate miRNAs as critical regulators of cancer development, which ultimately limits their current efficacy as routine clinical biomarkers. A recent analysis by Zhu et al. used the Cancer Genome Atlas (TCGA) database to establish nine miRNA targets of interest which were able to differentiate patients with cancer from controls (miR-125a, miR-125b, miR-129, miR-144, miR-194, miR-217, miR-328, miR-375, and miR-486) [14,15]. Unfortunately, none of these targets were evaluated in the current analysis. While commercially available multigene signatures, such as the OncotypeDX© 12 gene expression assay (Genomic Health Inc., Redwood City, CA, USA), have been incorporated into the paradigm for risk stratification in the setting of stage II/III colon cancer [70], efforts to discover novel biomarkers capable of surpassing CEA as a diagnostic biomarker for CRC have been futile. While our encouraging results suggest that miRNA may be an avenue worthy of further exportation, we must acknowledge that these efforts must be balanced with the challenges faced in successfully amplifying miRNA consistently across all biological tissues, which in essence limits the conclusions that may be drawn from the current analysis when considering the results in isolation.

Moreover, we must acknowledge that the current analysis suffers from several other limitations. Firstly, and most importantly, the failure to accurately differentiate median measurements in cancer patients and controls when amplifying these miRNAs in circulation limits the robustness of these results. The comprehensive amplification and further validation of these five miRNAs in the circulation is necessary before they can be further trialled or implemented in clinical practice to aid in CRC diagnostics. Secondly, this analysis incorporated samples and patient data from a single translational research centre where the recruited patients represent an isolated island population on the extreme of mainland Europe, leading to there being limited genetic diversity and relative homogeneity within this patient group [71,72]. Ultimately, research on the relevance of these biomarkers in international studies will be warranted prior to their use in clinical practice. Finally, in oncological patient treatment, it is now apparent that the novel taxonomy of CRC characterises it as a heterogenous disease composed of several distinct molecular subtypes [73]. Thus, the authors of this study may be seen by some to have wrongfully merged all 74 patients as one under the umbrella term ‘CRC’. Further analyses should be designed to have adequate power to provide insight into differences between cancer subtypes, if possible. However, in spite of these reported pitfalls, the authors wish to emphasise that the current observational study provides real-world evidence of the ability of the combination of miR-21, miR-31, and miR-150 in a five-miRNA oncogenic signature to aid diagnostics within the CRC paradigm.

In conclusion, our analysis identified and validated the differential expression profiles of miR-21, miR-31, and miR-150, which are capable of the substratification of cancerous and ‘normal’ tissues. When combined in a five-miRNA signature with two endogenous controls, these biomarkers have strong diagnostic accuracy, sensitivity, and specificity in identifying CRC tissue compared to TAN and CRC patients from ‘normal’ controls. While miR-21, miR-31, and miR-150 have previously been acknowledged as oncogenic miRNA within the context of CRC, to the best of our knowledge, this is a key study combining and amplifying these biomarkers within a diagnostic signature capable of identifying patients with CRC. Further experimentation with this novel five-miRNA signature is required in order to elucidate its relevance in amplifying circulatory tumour-associated miRNA, which may lead to the detection of CRC in even the earliest stages of the disease in the clinical setting.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/gidisord4030018/s1, Table S1: Details of the microRNA probes procured for this analysis, Figure S1: Receiver Operating Characteristic curves for (A) miR-31 and (B) miR-150 in differentiating tumour from tumour-associated normal tissues.
Author Contributions: All the authors were involved in the preparation of this manuscript. Study conceptualization was performed by M.G.D., G.F., N.M. and M.J.K.; M.G.D., M.P., E.H., A.J.L., N.M. and M.J.K. had access to raw data. N.M. had full access to all data and takes responsibility for the integrity of the data and the accuracy of the data analyses and had the final responsibility for the decision to submit for publication. M.G.D. and N.M. had the idea for the study. E.H. and M.P. performed statistical analyses. M.G.D., G.F. and H.A. performed the microRNA analysis. Analyses and interpretation of data were done by M.G.D. and N.M.; M.G.D. and N.M. drafted and wrote the manuscript, while all other authors were consulted with several drafts to appraise the intellectual content of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: Matthew G. Davey and Gerard Feeney received funding from the National Breast Cancer Research Institute, Ireland.

Institutional Review Board Statement: Ethical approval was obtained from the Galway Clinical Ethical Research Committee (Protocol Number 45/05 and CA.151). All methodology employed in this study followed the 1964 Helsinki Declaration and subsequent amendments or similar ethical standards. All experimental protocols were approved by the Department of Surgery at the National University of Ireland. Galway.

Informed Consent Statement: Informed consent was obtained from all participants for inclusion in the corresponding study.

Data Availability Statement: Date will be made available upon reasonable request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References


43. Slaby, O.; Srovoda, M.; Fabian, P.; Smerdova, T.; Knoflickova, D.; Bednarkova, M.; Nenutil, R.; Vyzula, R. Altered Expression of miR-21, miR-31, miR-143 and miR-145 Is Related to Clinicopathologic Features of Colorectal Cancer. Oncology 2007, 72, 397–402. [CrossRef]


45. Chang, K.H.; Mestdagh, P.; Vandesompele, J.; Kerin, M.J.; Miller, N. MicroRNA expression profiling to identify and validate reference genes for relative quantification in colorectal cancer. BMC Cancer 2010, 10, 175. [CrossRef]

46. Hellemans, J.; Mortier, G.; De Paepe, A.; Speleman, F.; Vandesompele, J. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. Genome Biol. 2007, 8, R19. [CrossRef]


50. Schee, K.; Boye, K.; Abrahamsen, T.W.; Fodstad, Ø.; Flatmark, K. Clinical relevance of microRNA miR-21, miR-31, miR-92a, miR-101, miR-10a and miR-145 in colorectal cancer. BMC Cancer 2012, 12, 505. [CrossRef]


