Deciphering the Functional Analysis of Differentially Expressed MicroRNAs Associated with Colorectal Cancer

Syarah Syamimi Mohamed  
Universiti Sains Malaysia

Azmir Ahmad  
International Islamic University Malaysia

Tengku Ahmad Damitri Al-Astani Tengku Din  
Universiti Sains Malaysia

Md Salzihan Md Salleh  
Universiti Sains Malaysia

Andee Dzulkarnaen Zakaria  
Universiti Sains Malaysia

Zaidi Zakaria (drzaidi@usm.my)  
Universiti Sains Malaysia

Research Article

Keywords: Colorectal cancer, microRNA profiles, mRNA target prediction, gene ontology and pathways interaction

Posted Date: March 13th, 2023

DOI: https://doi.org/10.21203/rs.3.rs-2606011/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License. Read Full License
Abstract

**Background:** Colorectal cancer (CRC) is a cancer that develops from clusters of abnormal crypt foci that grow in the lining of the colon or rectum. CRC develops into polyps before transforming into an adenocarcinoma. Evidences have suggested that enormous cancer cell proliferation is a result of gene dysregulation. For decades, researchers have discovered a link between microRNA (miRNA) and disease pathogenesis via post-transcription regulation of their target gene and alteration of protein translation. Despite of the evidences of a miRNA-CRC relationship with CRC pathogenesis, miRNA activities remain equivocal, and the target mechanism remains poorly characterized. Thus, we are attempting to decipher the functional assessments of differentially expressed miRNAs (DEMs) as well as their pathways interaction. We also aim to identify the list of miRNA/s which are highly associated with CRC.

**Results:** The DEMs were generated using microarray profiling of CRC tumor and non-tumor tissue. The DEMs were analysed using Gene Spring and a moderate T-test. The data is shown in a heatmap analysis and a volcano plot, with the filter set to $p$-value < 0.005 and fold change $\geq$ 2. We found 14 DEMs, 9 of which were upregulated and 5 of which were downregulated. The g:Profiler and Database for Annotation, Visualisation, and Integrated Discovery (DAVID) are used to identify the gene ontology and pathways to broaden the findings on their predicted function and pathway interaction. Our DEMs targeted the genes that are significantly associated with CRC pathogenesis, but we were unable to determine which miRNA specifically act as one. Therefore, we predicted the miRNAs function individually and we identified and suggested hsa-miR-20a-5p, hsa-miR-21-5p, hsa-miR-23a-3p, hsa-miR-24-3p and hsa-miR-29a-3p to play profound roles in CRC pathogenesis. Based on the evidences provided by web tools g:Profiler and DAVID, these miRNAs were discovered to commonly interact with PI3K-Akt signaling pathway, WNT signaling pathway and FGFR signaling.

**Conclusion:** To summarise, even though all of the DEMs were found to be significantly expressed in CRC tumoral tissue, it does not appear that they are strongly linked to CRC pathogenesis. Thus, with the data enriched by web tools analyses may aid in specifying potential miRNAs to be studied further. Apart from deciphering the miRNAs function and their interactive pathway, our finding also provides an additional knowledge on the specific potential miRNAs for future biomarkers of CRC.

Introduction

GLOBOCAN revealed that over the last two years, 10% of the global population has been diagnosed with colorectal cancer (CRC), with a 9.4% fatality rate from the total cancer-related death (1). This alarming trend is most likely owing to the disease's late development (2, 3) and being found in older people, making it difficult to diagnose and treat (4). Even though various methods have been developed to identify the condition, we still need a highly specific and sensitive screening tool. Molecular screening has been proposed to detect early-onset CRC, and various groups have attempted to profile significant miRNA dysregulated in CRC patients (5–7). However, the molecular mechanism of miRNA's targeting cancer causing-genes are still poorly understood.
MiRNA is a short nucleotide that alters or degrades the protein expression process after transcription. Canonically, the incorporation of mature miRNA with AGO protein, a component of RISC, will bind to the 3' UTR of target miRNA and initiate the silencing process by inhibiting protein synthesis via translational repression (8), deadenylation of mRNA and decay or transcript cleavage (9). MiRNA binds to the complementary region of mRNA due to sequence complementarity. A high specificity of sequence complementary promotes mRNA degradation, whereas an inadequate complementary inhibits protein translation (10).

MiRNAs can function as oncomiRs or tumor-suppressor genes (TS-Mirs). Surprisingly, the expression of both types of miRNAs, oncomiRs and TS-miRs, was shown to be synchronised in a negative feedback relationship. oncomiRs overexpression has been shown to repress TS-miRs. The similar thing happened when TS-miRs were overexpressed, which can lead to oncomiRs downregulation (11). Depending on the type of cancer and the impacted miRNA, inhibition or stimulation of these oncomiRs or TS-miRs may significantly reduce cancer cell proliferation (12).

Previous research managed to highlight the specificity of DEMs in CRC tissue and non-tumoral adjacent tissue (13–15). Regardless of data robustness, none can be promised to contribute for CRC therapy due to its divergence of mechanism. In addition, the pattern of miRNA may be difference according to race and region (16) and less were reported miRNA dysregulation in Asian population specifically for CRC. Therefore, the present study aims to establish the preliminary profile of a miRNAs from the Asian Malay population and stratify the DEMs based on disease association, target genes, gene ontology, and interacted pathway.

Methods

Sample collection

During a surgical operation at Hospital Universiti Sains Malaysia (HUSM), 12 pairs of CRC tissue (tumor and non-tumor) were obtained (Kelantan, Malaysia). The entire recruitment approach meets the inclusion requirements for CRC. Patients aged 30 to 70 were chosen, and their ethical consent was authorised by The Human Research Ethics Committee of USM (JEPeM), USM: USM/JEPeM/18050239. In order to validate the tumoral tissue, histological data was gathered.

Total RNA processing and miRNA profiling.

Following the GeneAll Hybrid-R kit’s instructions, RNA was taken from 12 pairs of fresh-frozen mucosa regions from both tumour and non-tumor tissue. The glass beads method was used to mix up about 80–100 mg of colonic mucosa tissue. Before doing the microarray profiling analysis, the final amount of RNA was measured with a Bioanalyzer for RNA integrity number (RIN). Before being labelled with Cyanine 3-pCp for hybridization, all of the RNA samples were normalised to 50ng using the miRNA Complete Labelling and Hyb kit (Agilent, Valencia, CA). Later, this RNA is cleaned up by using a vacuum concentrator to dry it. The labelled miRNA was then mixed with 10 Hi-RPM Hybridization buffer, 2 Hi-RPM
Hybridization spike-in solution, and 2 Agilent blocking agent. After incubation at 100°C, samples are immediately hybridised to Agilent SurePrint G3 Human miRNA Microarray 860K (Design ID: 070156) for 20 hours at 55°C in a rotating Agilent microarray hybridization oven. After the chips were hybridised, they were washed for 5 minutes with GE Wash Buffer 1 (Agilent) at room temperature and 5 minutes with GE Wash buffer 2 at 37°C (Agilent). Chips were scanned right away using an Agilent SureScan Microarray Scanner (G4900DA) with an extended dynamic range (10–100%) and a 3m resolution at a wavelength of 532nm.

**Identifying differently expressed miRNAs (DEMs)**

Agilent Gene Spring Analysis Software version 14.9.1 is used to identify the differently expressed miRNAs (DEMs) between the type of tissue. Moderate- T-test is used with cut off values of \( p \)-value 0.05 and fold change 2.0 to consider it statistically significant, whereas fold change 2 is regarded to be upregulated miRNAs and fold change – 2 is considered to be downregulated miRNAs.

**MiRNA-mRNA prediction for functional analysis**

Association of each DEMs with diseases is retrieved from miRCancer databases (http://mircancer.ecu.edu/). The validated target gene numbers are filtered using miRWalk (http://mirwalk.umm.uni-heidelberg.de/), an open-source platform that incorporated with others databases such as TargetScan, miRDB and miRTarBase (17). Next, we specified the target genes that are associated to CRC using DIANA micro T. This database is also freely available on http://www.microrna.gr/webServer. The predicted genes obtained from DIANA micro T are the genes that related to KEGG pathways that matches for CRC, where the threshold level is set to 0.7/0.8.

**Gene ontology (GO) and KEGG pathway analysis of DEMs.**

The Gene Ontology (GO) functional enrichment analysis of target genes was performed using g:Profiler, https://biit.cs.ut.ee/gprofiler/gost and database for annotation, visualization and integrated discovery (DAVID), https://david.ncifcrf.gov/ that provides the gene ontology enrichment on molecular function, cellular component and biological process. It also provides the biological pathways from three different sources as Kyoto Encyclopedia of Genes and Genomes (KEGG), Reactome and Wikipathways.

**Results**

**Demographical and clinicopathological data**

Table 1 showed the clinicopathological data from the samples recruited. The average of recruited CRC patients was 64.7 ± 14.3 years with 41.7% male and 58.3% female. Their fresh colorectal tumor tissue and non-tumoral adjacent tissue were collected during surgical procedure. Based on histopathological analysis, none were recruited with stage , 25% of the cases are stage , 41.7% are in stage and 33.3% in stage . Most of the CRC phenotype is rectosigmoid (66.7%). Meanwhile rectal and caecum phenotype
both with 16.7% respectively. Only 2 samples were seen in mucinous and the rest are adenocarcinoma. Figure 1 shown the Hematoxylin and Eosin (H&E) staining of tissue phenotypes.

Table 1
Demographical and clinicopathological data of recruited samples

<table>
<thead>
<tr>
<th>Variable</th>
<th>CRC tissue, n = 12 (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>5.0 (41.7)</td>
</tr>
<tr>
<td>Female</td>
<td>7.0 (58.3)</td>
</tr>
<tr>
<td>Age</td>
<td></td>
</tr>
<tr>
<td>Median age</td>
<td>68.5</td>
</tr>
<tr>
<td>Mean ± standard deviation</td>
<td>64.7 ± 14.3</td>
</tr>
<tr>
<td>TNM Stage</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td></td>
</tr>
<tr>
<td>3.0 (25.0)</td>
<td></td>
</tr>
<tr>
<td>5.0 (41.7)</td>
<td></td>
</tr>
<tr>
<td>4.0 (33.3)</td>
<td></td>
</tr>
<tr>
<td>Metastasis</td>
<td></td>
</tr>
<tr>
<td>Liver/Lung</td>
<td>6.0 (50.0)</td>
</tr>
<tr>
<td>Others</td>
<td>-</td>
</tr>
<tr>
<td>None</td>
<td>6.0 (50.0)</td>
</tr>
<tr>
<td>Tumor lesion</td>
<td></td>
</tr>
<tr>
<td>Rectosigmoid</td>
<td>8.0 (66.7)</td>
</tr>
<tr>
<td>Rectal</td>
<td>2.0 (16.7)</td>
</tr>
<tr>
<td>Caecum</td>
<td>2.0 (16.7)</td>
</tr>
<tr>
<td>Histology</td>
<td></td>
</tr>
<tr>
<td>Adenocarcinoma</td>
<td>10.0 (83.3)</td>
</tr>
<tr>
<td>Mucous</td>
<td>2.0 (16.7)</td>
</tr>
</tbody>
</table>
Identification of DEMs.

The Gene Spring analysis tool was used to screen the differential expressed miRNAs (DEM) using \( p \)-value < 0.05 and fold change \( \geq 2 \) as the cut-off criteria. The list of significant DEMs were presented in heatmap and a volcano plot (Fig. 2, A-B). Quantitatively, these DEMs are presented in a bar graph as in Fig. 3. 14 significant DEMs were profiled, including 9 upregulated miRNAs (hsa-miR-192-5p, hsa-miR-20a-5p, hsa-miR-21-5p, hsa-miR-210-3p, hsa-miR-215-5p, hsa-miR-223-3p, hsa-miR-23a-3p, hsa-miR-24-3p and hsa-miR-29a-3p) and 5 downregulated miRNAs (hsa-miR-3653-3p, hsa-miR-3945, hsa-miR-5090, hsa-miR-5196-5p and hsa-miR-6133) as shown in Fig. 3.

MiRNA-mRNA prediction for functional analysis

We provide the target genes data obtained from miRWalk and DIANA micro T web tools in Additional file 1.

GO analysis and pathway

Based on the g:Profiler analysis, we found “aspartic-type endopeptidase activity” and “apeptidase activity” to be the significantly enriched molecular function. Notably, “phosphatidylserine exposure on apoptotic cell surface database” is the only one biological processes enriched, while “late endosome lumen” and “endosome lumen” was found to be related to the cellular component as can be seen in Fig. 4. Conversely, the DAVID presented more data on the GO and pathway. We listed the top 10 significant data for each of them as listed in Table 2. These identifications suggest the potential functions of our input genes which were based on our profiled DEMs.
## Table 2
The top 10 of significant GO and pathway retrieved from DAVID tool.

<table>
<thead>
<tr>
<th>Gene ontology (p-value)</th>
<th>Pathway (p-value)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biological process</td>
<td>Cellular component</td>
</tr>
<tr>
<td>GO:0035556 Intracellular signal transduction (0.0045)</td>
<td>GO:0032991 Macromolecular complex (0.0017)</td>
</tr>
<tr>
<td>GO:0021772 Olfactory bulb development (0.0080)</td>
<td>GO:0005654 Nucleoplasm (0.0018)</td>
</tr>
<tr>
<td>GO:0007265 Ras protein signal transduction (0.0092)</td>
<td>GO:0016020 Membrane (0.0029)</td>
</tr>
<tr>
<td>GO:1903214 Regulation of protein targeting to mitochondrion (0.0102)</td>
<td>GO:0097486 Multivesicular body lumen (0.0076)</td>
</tr>
<tr>
<td>GO:0070782 Phosphatidylserine exposure on apoptotic cell surface (0.0102)</td>
<td>GO:0051233 Spindle midzone (0.0126)</td>
</tr>
<tr>
<td>GO:0007264 Small GTPase mediated signal transduction (0.0107)</td>
<td>GO:0005829 Cytosol (0.0132)</td>
</tr>
<tr>
<td>GO:0008286 Insulin receptor signaling pathway (0.0107)</td>
<td>GO:0048471 Perinuclear region of cytoplasm (0.0174)</td>
</tr>
</tbody>
</table>
### Gene ontology (p-value)

<table>
<thead>
<tr>
<th>Biological process</th>
<th>Cellular component</th>
<th>Molecular function</th>
<th>KEGG</th>
<th>Reactome</th>
</tr>
</thead>
<tbody>
<tr>
<td>GO:0050731</td>
<td>GO:0005885</td>
<td>GO:0030515</td>
<td>hsa05164</td>
<td>R-HSA-9006936</td>
</tr>
<tr>
<td>Positive regulation of peptidyl-tyrosine phosphorylation (0.0118)</td>
<td>Arp2/3 protein complex (0.0189)</td>
<td>SnoRNA binding (0.0581)</td>
<td>Influenza A (0.0363)</td>
<td>Signaling by TGFB family members (0.0221)</td>
</tr>
<tr>
<td>GO:0007268</td>
<td>GO:0030426</td>
<td>GO:0016887</td>
<td>hsa04151</td>
<td>R-HSA-69416</td>
</tr>
<tr>
<td>Chemical synaptic transmission (0.0129)</td>
<td>Growth cone (0.0292)</td>
<td>ATPase activity (0.0668)</td>
<td>PI3K-Akt signaling pathway (0.0509)</td>
<td>Dimerization of procaspase-8 (0.0251)</td>
</tr>
<tr>
<td>GO:0001935</td>
<td>GO:0045202</td>
<td>GO:0019901</td>
<td>hsa01524</td>
<td>R-HSA-3371378</td>
</tr>
<tr>
<td>Endothelial cell proliferation (0.014508)</td>
<td>Synapse (0.0327)</td>
<td>Protein kinase binding (0.0670)</td>
<td>Platinum drug resistance (0.0784)</td>
<td>Regulation by c-FLIP (0.0251)</td>
</tr>
</tbody>
</table>

We also identified the GO and pathway involved based on individual DEMs presented using g:Profiler database and we supply the findings in Supplementary data 2. Each of the miRNA coded differently for their GO and pathway involved. However, g:Profiler is unable to provide the information on hsa-miR-210-3p, hsa-miR-215-2p, hsa-223-3p, hsa-miR-3945, hsa-miR-5090, hsa-miR-5196-5p, and hsa-miR-6133. Based on GO analysis, it is revealed that various function in molecular function, cellular component and biological process for each of miRNAs. Their respective biological pathways were also obtained from three different databases, KEGG, REACTOME and WikiPathways, provided in Additional file 2.

### Discussion

Corroborative evidences have identified the potential use of miRNAs as genomic biomarker for disease diagnosis, specifically for CRC (18–20). However, there is still limited consensus on a single or a group of miRNAs that can solidly marked as CRC biomarker. The role of MiRNAs are heterogenous, they can act both as oncoMir by promoting cell growth or TS-miRs that control the cell growth (11). Thus, identifying the role of DEMs in CRC patients and associated signatures is beneficial in increasing our understanding of these molecular types. We screened the DEMs and investigated their functions and mechanisms. Based on web tool evaluations, we attempted to determine which miRNAs were strongly related to CRC.

To answer the association of miRNAs with potential related disease, a data from miRCancer database was retrieved. According to analyses, only hsa-miR-192-5p, has-miR-20a-5p, hsa-miR-21-5p, hsa-miR-215-5p, hsa-miR-24-3p, and hsa-miR-29a-3p were reported to be related with CRC (see Additional file 1). Although these miRNAs seem to be potential biomarker for CRC, we also discovered that each of the
DEMs were also linked to other disease, such as hepatocellular carcinoma, lung adenocarcinoma, gastric cancer, osteosarcoma, rectal cancer, glioblastoma, ovary cancer, esophageal squamous cell carcinoma, and other. The dynamic results of the identified miRNAs are possibly due to the capability 2–8 of its 5’ complement sequence of the same miRNA that binds to several different of 3’ untranslated region (UTR) of mRNA (10, 21). For example, a study by Mokhlesi and Talkhabi concluded the hsa-miR-192-5p could potentially target genes associated with lung adenocarcinoma (22), but two years later, another study by Toolabi and team predict that hsa-miR-192-5p may targets genes associated with colon cancer (23).

According to our clinicopathological data in Table 1, half of the recruited samples have liver/lung metastases. We hypothesised that the CRC tumor in our patients had metastasized to other areas of the body, or that it was a secondary tumor that was spread from another part of the body, such as the lung or liver. Due to lack of relevant data, we were unsure to explain the relationship of the given DEMs with other disorders, particularly in lung and liver cancer. However, previous study claims that immunohistological examination can be perform to distinguish whether the metastasis originates from the colon or from other sources (24). Thus, from a clinical standpoint, patients diagnosed with CRC should be tested for other types of cancer, and the tumor excised should be subjected to immunohistological investigation (25). Histologically, the mucinous type of tissue might have different dysregulation of miRNAs from adenocarcinoma type of tissue. Mucinous histology is frequently correlated with the alteration of BRAF, PIK3CA and growth β-factor pathway than non-mucinous histology (26). In our study, we managed to have a minor number of mucinous tissue type, therefore we unable to analyse DEMs histologically.

The divergence of miRNA expression is also a point of contention. According to one study, hsa-miR-192 is downregulated in colon adenoma tissue (27), which contradicts our profile result of upregulation. There are various factors that can alter the dynamical expression of miRNAs such as miRNA availability and abundance in distinct cell types, miRNA compartmentalization, or cell state way (28). Interestingly, other researchers discovered that several of these miRNAs (hsa-miR-20a-5p, hsa-miR-21-5p and hsa-miR-24-3p) were also present in non-tumor tissue samples (29, 30). The presence of miRNAs in other sample sources, such as feces, urine and blood, may explained by miRNA transportation (31–33) and its ability to communicate with the nearby cells (34–36). Their conserved (37, 38) and stable features (12, 39, 40) make them as a convenient screening tools for disease diagnosis in future use.

Even so, there are few miRNAs that might not be reported yet in miRCancer database, as the results remains elusive or unavailable, such as hsa-miR-3653-3p, hsa-miR-3945, hsa-miR-5090, hsa-miR-5196-5p, and hsa-miR-6133. Notably, miRWalk database also unable to retrieve the target genes for these downregulated DEMs. However, the DIANA-micro T database can identify the target genes for the significantly expressed miRNAs, including the downregulated miRNAs identified in our work (see Additional file 1). Since there are numerous of available databases in identifying miRNAs mechanisms, the choice of databases is crucial as they pertained different algorithms. Despite of the different capabilities of each prediction tools, the prediction still run based on the common features of miRNA which are seed-match, conservation, free energy and site complementarity, and accessibility (41).
We enriched the miRNAs’ function by identifying their GO and pathway. We used two enrichment predictive tools; g:Profiler as shown by the result in Fig. 4 and DAVID in Table 2. The g:Profiler database is a free web tools that equipped with informative visual of results presentation. Its features that support the ranked gene lists reveal the potential GO based on the significant values from our listed gene. As presented in Fig. 4, the genes targeted by our DEMs have “aspartic-type endopeptidase” and “apeptidase activity” as the molecular function (MF). The biological process (BP) is enriched with “phosphatidylserine exposure on apoptotic cell surface” only but for the cellular component (CC), the genes are encoded for “late endosome lumen” and “endosome lumen”.

Both “aspartic-type endopeptidase” and “apeptidase activity” were classified under peptidase activity. High expression on peptidase activity lead to tumor invasion, angiogenesis and metastasis. In addition, peptidases are crucial to modulate apoptosis, regulate immune response for anti-tumor, and promote growth and development of cancer stem cells and the transition of epithelial to mesenchymal cell type (42). Prior this, there is study reported that the “aspartic-type endopeptidase” is enrich in biological process of gastric cancer tissue (43). Similarly, the genes mostly contribute for “phosphatidylserine exposure on apoptotic cell surface” for BP that has been reported to be induced by apoptotic factor (44) and mostly associated with endothelial-tumor cells (45). For CC, our targeted genes are active in “late endosome lumen” and “endosome lumen”, which actively involved in transportation of intracellular membrane vesicles (46). In short, our dynamics GO analyses may be heavily reliant on the clinicopathological data that we have collected, where half of samples were diagnosed with advanced stage and the cancer cell have met to liver/lung mostly. Thus we postulated that our GO analysis revealed from g:Profiler, highly associated with tumorigenesis within the lumen lesion and cell apoptosis.

On the other side, the GO and targeted pathway predicted by DAVID are more robust. We ranked the data as can be seen in Table 2. DAVID also reveal the same enriched BP as g:Profiler which is phosphatidylserine exposure on apoptotic cell surface as one of the presented top 10 data. Again, the DAVID is synchronized with g:Profiler when reported “multivesicular body lumen” for the CC and “aspartic-type endopeptidase activity” for the MF. DAVID has more privilege in predicting the targeted pathway. We also managed to retrieved the additional pathways from two databases through DAVID; KEGG and Reactome. KEGG mostly reported on cell signaling pathway. E.g. “MAPK signaling pathway”, “Ras signaling pathway”, “Neurotrophin signaling pathway”, “Rap1 signaling pathway”, “PI3K-Akt signaling pathway”. Meanwhile, Reactome revealed the transcriptional activity. E.g. “Downregulation of SMAD2/3:SMAD4 transcriptional activity”, “Signal transduction”, “Transcriptional Regulation by TP53”, “Dimerization of procaspase-8”, and “Regulation by c-FLIP”.

Aside from presenting the data on targeted genes and ranking probable mechanisms from the DEMs given, we also examined the DEMs separately, which is expected to aid future decisions in researching a panel of single miRNAs or clusters of miRNAs as CRC biomarkers. According to individual analyses of DEMs, as shown in Additional file 2, we suggested that hsa-miR-20a-5p, hsa-miR-21-5p, hsa-miR-23a-3p, hsa-miR-24-3p, and hsa-miR-29a-3p either as a panel/panel cluster for CRC potential biomarkers based on the pathway recognition. They were recognized to commonly interact with “PI3K-Akt signaling
pathway”, “WNT signaling pathway” and “FGFR signaling” that is highly associated with CRC pathogenesis.

Despite of promising evidence from web tools used, our enrichment finding analyses might be different from evidences provided through Malacards, The Human Disease Database, (https://www.malacards.org/#), which is an integrated database that compiled the information on genomic-associated disease. According to Malacards, the top 10 pathway that are directly related to CRC are signaling by “TCF7L2 Mutants”, “Defective MUTYH Substrate Processing”, “Defective MUTYH Substrate Binding” and “Defective Base Excision Repair Associated with MUYTH”. Other pathways that associated with CRC are “ERK Signaling”, “Signal Transduction”, “TGF-Beta Pathway”, “Prolacting Signaling”, ”RAF/MAP Kinase Cascade”, “Phospholipase-C Pathway”, “Beta-Adrenergic Signaling”, “Signaling by WNT” and “NFAT and Cardiac Hypertrophy”.

Collectively, we believe that our data provide additional knowledge and insight regarding miRNAs underlying mechanisms in CRC of Asian Malay population. Apart from the genetic make-up, the study of miRNAs across the specific population might need to consider the gender, environmental causal and demographical barriers of population as well. For the time being, less report is critically discussed on this evidence. Although our findings showed that not all of the screened DEMs are linked to CRC, we believed that this could contributed to the miRNA’s profiling data, especially in Asian Malay community. Prior this, CRC was mostly diagnosed in western countries (47, 48) and recently was identified in Asian and Gulf country (49, 50). Seeing this emerging cases, a more practical approach in identifying epigenetic miRNAs as CRC biomarker is needed. Although our study may seem lack of data in comparisons of DEMs across the populations, we believe that this study provides an insight on miRNAs expression in Asian Malay population since less report is found regarding this. Other screening based on sex or stage is also preferable for possible correlation with clinicopathological data, because miRNA expression is epigenetic and may be regulated by other environmental factors.

**Conclusion**

This study is applicable in the pursuit of understanding the mechanism of DEMs in CRC tumoral tissue by providing information from disease association, targeted mRNAs, gene ontology, and the pathways interacted. These findings can be extended to therapeutic miRNA targets via their respective pathways. Those identified miRNAs will need to be validated on a larger sample size with a specific target in order to gain a better understanding of their mechanism in CRC pathogenesis.

**Abbreviations**

Declarations

Ethics approval and consent to participate

This study was approved by The Human Research Ethics Committee of USM (JEPeM), USM: USM/JEPeM/18050239, and written informed consents were obtained by patients themselves and their relatives, voluntarily. All procedures are carried out in accordance with the Helsinki Declaration's relevant guidelines and regulations under 'Informed Consent' section. Prior to providing informed consent, participants were briefed on the study's objective and outcome. Researchers have consulted with clinicians to determine the impact of this study on patients' personal conditions. Personal information about those involved was kept private.

Consent for publication

Not applicable

Availability of data and interest

All data generated or analysed during this study are included in this published article [and its supplementary information files].

Competing interest

Authors declared that they have no competing interests.

Funding

This study is fully funded by Research University Grant from Universiti Sains Malaysia, RUI 1001/PPSP/8012260.

Authors’ contributions

Zaidi Zakaria and Tengku Ahmad Damitri Al Astani Tengku Din designed the project and revised the research and manuscript critically. Andee Dzulkarnaen Zakaria and Md Salzihan Md Salleh contributed to clinical expertise and critically revised the research analyses. Syarah Syamimi Mohamed performed the experiments and wrote the manuscript and Azmir Ahmad made a contribution to data analysis and manuscript writing. All authors read and approved the final manuscript.

Acknowledgement

This study was supported by grants from university, Universiti Sains Malaysia (RUI 1001/PPSP/8012260) and all study was taken place in university. Authors would like to acknowledge experts from International Islamic University of Malaysia, National University of Malaysia and National Institute of Health, Malaysia for willingly to share ideas and critiques on research without wanting to consider for any interest.
References

12. Svoronos AA, Engelman DM, Slack FJ. OncomiR or Tumor Suppressor? The Duplicity of MicroRNAs in Cancer. 2016; Available from: www.aacrjournals.org


Figures
Figure 1

(A) is the H&E sections showed the tumor cells floating in large extracellular lakes if mucin for mucinous type of sample. Meanwhile (B) showed the infiltration of tumor beyond the mucularis propria and into the pericolic fat. The tumor invasion of perineural and lymphovascular were seen for moderately differentiated adenocarcinoma samples.
Figure 2

(A) The heatmap depicted the expression of DEMs in CRC tumor tissue compared with their non-tumor adjacent tissue. Heatmap of 12 paired mucosa tissue presented 14 significantly DEMs, with 9 were upregulated (green box colored) and 5 were downregulated (red box colored). On the other side, (B) The volcano plot that filter all the DEMs and presented 14 significant DEMs that passed the fold change $\geq 2$ and $p$-value < 0.05.
Figure 3

The quantitative bar graph for DEMs that passed the set up cut-off value, p<0.05 and fold change ≥ 2.

Figure 4

The functional analysis of targeted genes was shown in a colored bar graph, with red dots representing molecular function, yellow dots representing biological processes, and green dots representing cellular components. These gene ontologies are highlighted because they have a significance p-value.

Supplementary Files
This is a list of supplementary files associated with this preprint. Click to download.

- Additionalfile1.docx
- Additionalfile2.docx