



Article

Functional Prediction of Candidate MicroRNAs for CRC Management Using in Silico Approach

Adewale Oluwaseun Fadaka *, Ashley Pretorius and Ashwil Klein

Department of Biotechnology, Faculty of Natural Sciences, University of the Western Cape, Private Bag X17, Bellville, Cape Town 7535, South Africa; aspretorius@uwc.ac.za (A.P.); aklein@uwc.ac.za (A.K.)

* Correspondence: afadaka@uwc.ac.za; Tel.: +27-63-051-1928/+23-48-039-2420-52

Received: 29 August 2019; Accepted: 14 September 2019; Published: 19 October 2019



Abstract: Approximately 30–50% of malignant growths can be prevented by avoiding risk factors and implementing evidence-based strategies. Colorectal cancer (CRC) accounted for the second most common cancer and the third most common cause of cancer death worldwide. This cancer subtype can be reduced by early detection and patients' management. In this study, the functional roles of the identified microRNAs were determined using an in silico pipeline. Five microRNAs identified using an in silico approach alongside their seven target genes from our previous study were used as datasets in this study. Furthermore, the secondary structure and the thermodynamic energies of the microRNAs were revealed by Mfold algorithm. The triplex binding ability of the oligonucleotide with the target promoters were analyzed by Trident. Finally, evolutionary stage-specific somatic events and co-expression analysis of the target genes in CRC were analyzed by SEECancer and GeneMANIA plugin in Cytoscape. Four of the five microRNAs have the potential to form more than one secondary structure. The ranges of the observed/expected ratio of CpG dinucleotides of these genes range from 0.60 to 1.22. Three of the candidate microRNA were capable of forming multiple triplexes along with three of the target mRNAs. Four of the total targets were involved in either early or metastatic stage-specific events while three other genes were either a product of antecedent or subsequent events of the four genes implicated in CRC. The secondary structure of the candidate microRNAs can be used to explain the different degrees of genetic regulation in CRC due to their conformational role to modulate target interaction. Furthermore, due to the regulation of important genes in the CRC pathway and the enrichment of the microRNA with triplex binding sites, they may be a useful diagnostic biomarker for the disease subtype.

Keywords: microRNA; CRC; functional predictions; in silico analysis; triplex binding site

1. Introduction

With increasing incidence and mortality of cancers, colorectal cancer (CRC) remains one of the leading malignant cancers [1]. Currently, tumor resection followed by chemotherapy is one of the most effective treatments for CRC but recurrences are inevitable [2]. Approximately 45% of recurrences are in the first year after initial tumour removal, and over 90% of recurrences occur within four years [3]. In CRC, staging and pathological characteristics are one of the major predictors of diagnosis to enhance treatment options [4]. Thus far, there are limited biomarkers to predict CRC. The search for more biomarkers that are specific for staging is required.

Recent research suggests that microRNAs regulate many gene functions in human cancers [5], and these oligonucleotide sequences have been proposed as novel biomarkers for cancers [6]. Previous studies have shown that their expressions are altered in numerous types of cancers including CRC [7,8]. They also act as regulators of gene expression [9] through gene repression and/or mRNA degradation in many biological processes such as apoptosis, cell development, cell differentiation, cell proliferation,

and metabolism [10–12]. These biological processes are crucial in carcinogenesis [13]. Although several studies have reported an association between microRNAs and CRC development [14–17], the role of microRNAs forming triplexes with their targets to infer function in CRC has been barely explored. Several microRNAs have been reported to date but their laboratory functional determination in relation to CRC is challenging and time-consuming.

In silico approaches have been developed to connect sequences of microRNAs and their targets to infer function in Cancer studies. Biologists are mainly particular about the structural and functional properties of any newly derived sequence (protein or nucleotide). In silico predictions are therefore important for this discovery based on successful knowledge-based principles. These principles rely on the fact that the best way to predict the structure and/or function is to find similar sequences in existing databases using the information about them to infer conclusions about properties of the new sequence. Algorithms have been developed and implemented as computer programs (local or web-based tools) to perform this function.

The genomic sequence of many higher eukaryotes is now complete, and the patterns of expression of thousands of genes under diverse conditions are known. This offers researchers the opportunity to identify and analyze the parts of a genome believed to be responsible for most transcription control, known as the promoters [18]. Short nucleotide sequences, most especially non-coding RNAs, have been reportedly used in the induction of DNA cleavage at specific sites with the aid of triplex formation between the RNA and the DNA [19,20]. The various triplex formation between these nucleotides has suggested a role in gene expression regulation [21,22]. The binding of DNA leading to the triplex bond formation through a hydrogen bond to the third strand has been experimentally discovered but the biological functions are still unknown [23]. The interaction of microRNAs with DNA have not been studied extensively [24,25]. Reports also suggested that the interactions between microRNAs and gene promoter regions may play a more direct role in regulating the efficiency of transcription of the gene involved [26–28].

One of the roles includes the methylation of CpG islands [29]. However, other mechanisms yet unidentified or not fully studied could be possible for the direct interaction of microRNAs with genes for transcription regulation. Since dsDNA is capable of forming triplex structures through interactions with DNA or RNA in the major groove of the DNA duplex, researchers postulated that microRNA may form triplex structures with duplex DNA through either Hoogsteen or reverse Hoogsteen hydrogen bonds, and thereby directly interacting with target DNA sequences in regulatory regions including gene promoters in the human genome, with the potential to alter gene function [30–32]. Bioinformatics approaches revealed that the mammalian genome is composed of several triplex formation binding sites [24,30,33,34]. This suggests that tethering RNA to specific genomic sites might guide RNA-associated regulatory proteins to establish an epigenetic landscape that facilitates or inhibits gene expression. In this study, the function of the candidate microRNAs was predicted for CRC management using an in silico approach. Given the ever-expanding number of microRNAs, understanding their functional aspects through sequences represents a promising research area.

2. Results

Secondary Structure of the Candidate MicroRNAs

The function of a given microRNA molecule may be determined by sequence or structure that it is most likely to fold into. It may also be governed by whether small sequence have the ability to fold into a particular substructure even if the substructure does not appear in any optimal fold.

3. Discussion

Genes implicated as targets for the candidate microRNAs have been experimentally validated and their associations with CRC confirmed by previous studies [35–43]. Also, research has confirmed that the regulators of these targets have been experimentally determined and validated in other cancer

subtypes but not in CRC [44–54]. Therefore, the interaction of both the microRNAs and targets could suggest their involvement in CRC. The prognostic, diagnostic, and expression analysis performed on these genes together with the gene ontology analysis from our previous study suggest that they could be used as good diagnostic and therapeutic biomarkers in the detection and management of CRC.

3.1. Secondary Structure and the Thermodynamic Energies of the Candidate MicroRNAs

MicroRNAs have intrinsic potential to form secondary and tertiary structures by folding through base pairing. This folding is given primarily by the fact that certain nucleotides have an affinity for binding another nucleotide, for example (G-C and A-U/). Each nucleotide can only bind with one partner at a time and long stretches of rich nucleotide as a binding partner elsewhere on the sequence are favoured over short stretches or individual non-sequential pairing. These pairings compete with each other since any G in the sequence could pair with any C and similarly, any A with any U. The arrangement in Figure 1 may be optimal in terms of maximizing the length of the concurrent run and pair of nucleotides. These arrangements are only probabilistic and outside forces such as temperature and pH may induce a different pairing arrangement. The molecule may even transition randomly between numerous newly optimal structures based on the fluctuation in available energy. The mature sequences of the five candidate microRNAs were subjected to secondary structure prediction by Mfold and inspected manually against filtering criteria, as indicated by default to check for any discrepancies.

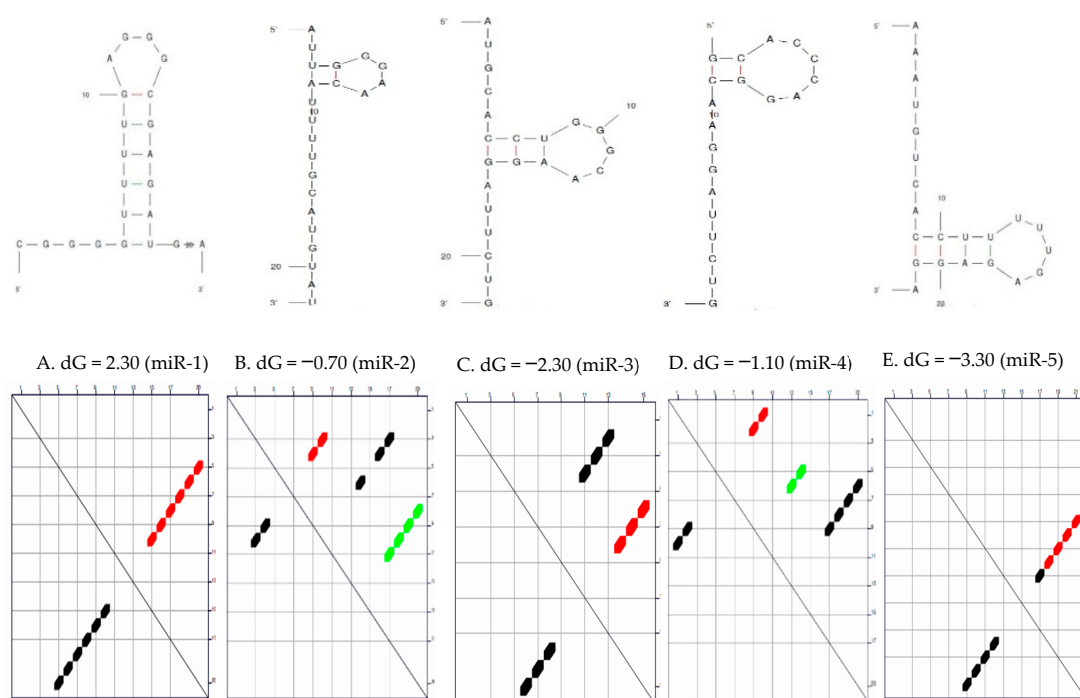


Figure 1. The predicted secondary structure of the five candidate microRNAs with their dot plot directly below showing the optimal energy. (A). microRNA miR-1 with its optimal energy of -2.30 ; (B). MicroRNA miR-2 with its optimal energy of -0.70 ; (C). microRNA miR-3 with its optimal energy of -2.30 ; (D). microRNA miR-4 with its optimal energy of -1.10 ; (E). microRNA miR-5 with its optimal energy of -3.30 . For the dot plots, red, black, and green dots represent all the optimal foldings (superposition of all possible sub-optimal foldings) Therefore, each colour represents a potential folding configuration.

The prediction of microRNA secondary structure is a long-established problem of computational biology which has received a lot of attention in recent years due to mounting evidence that underscores the significant of microRNA structure in a wide variety of biological processes [55–57]. The results showed that four of the five have the potential to form more than one secondary structure (folding

configuration). Studies have reported numerous microRNAs as biomarkers of disease and potential therapeutic targets due to the fact that their secondary structures may give insight into orchestrated microRNA-dependent gene regulation and be a step forward to understand their functions and involvement in carcinogenesis, and improve therapeutic designs. The results obtained from the secondary structure prediction of the candidate microRNAs are consistent with previous in silico structural determination studies and have indicated that over 70% of human microRNAs may fold into a hairpin structure and almost 70% could potentially form self-aggregated homoduplexes [58–60]. All the predicted microRNAs showed hairpin structure which could be enhanced by the cellular environment.

Valid prediction of microRNA–mRNA binding energies is crucial for the understanding of interactions. The thermodynamics of microRNA interactions can be understood as the sum of the energy necessary to ‘open’ the binding site and the energy gained from hybridization [57]. The thermodynamic parameters of the secondary structure of the candidate microRNAs are represented in Table 1 above. Ronchieri et al. [61] reported that microRNA secondary structures contribute to target recognition because there is an energetic cost to freeing base-pairing interactions within mRNA in order to make the target accessible for microRNA binding [62–64]. The ΔG values of the candidate microRNAs were calculated using Mfold (Figure 1 and Table 1). The optimal sequences were predicted to have optimal folding in miR-1 and 5, while the sub-optimal sequences were predicted to have sub-optimal folding in miR-2 and 4. Understanding the optimal folding of base pairs is the least likely secondary structure formation during the reaction [65]. The lower the free energy the more stability of the microRNA. From Table 1, miR-5 is the most stable microRNA, followed by miR-3 and miR-1, respectively.

Table 1. Parameters and secondary structures of the microRNA sequences.

S/N	MicroRNAs	Length	δG (kcal/mol)	Initial ΔG (kcal/mol)	Stru ^C
1	miR-1	22	0.0	–2.30 –0.70	1
2	miR-2	22	0.7	–0.40 0.00	3
3	miR-3	22	0.7	–2.80 –2.10 –1.10	2
4	miR-4	20	0.8	–0.80 –0.30	3
5	miR-5	22	0.6	–3.30 –2.70	2

δG : Free energy in plot profile; ΔG : Optimal energy of secondary structures (kcal/mol) at 37 °C with optimal and sub-optimal structures, respectively. Stru^C: Number of secondary structure calculated by Mfold.

3.2. CpG Island of the Promoter Sequences

The phenomenon of tumor alteration through epigenetic silencing associated with dense hypermethylation of CpG islands, and their complex interplay with modifications in histone structure, provides an alternate mechanism to genetic inactivation of tumor suppressor genes via loss or mutation [66]. Binding sites in the genome have a great regulatory impact on the gene activities in their neighborhood. Predictive tools are therefore essential for deciphering the overall regulatory potential of gene control regions such as promoters and enhancers. The CpG island analysis of the prioritized microRNA target genes (7 genes) is shown in Table 2 with Min. GC% of 51% observed in APC while Max. GC% of 83% was observed in KRAS. The ranges of the observed/expected ratio of CpG dinucleotides of these genes are from 0.60 to 1.22. CpG islands have been reported to be found in approximately 50% of human promoters [67,68]. Promoters with CpG islands are associated with housekeeping genes [69,70] and are identified by three primary characteristics: (1) they are over 200 base pairs in length, (2) they have over 50% GC composition, and (3) they retain an observed/expected ratio of CpG dinucleotides greater than 0.6 [71]. In normal tissues, CpG islands associated with tumor

suppressor genes are unmethylated but during tumor formation, they are often methylated. Evidence suggests that *de novo* methylation of CpG islands induces the silencing of associated tumor suppressor genes and may, in fact, be an important step in tumor formation [72,73]. The particular genes that are hypermethylated in tumor cells are strongly specific to the tissue of origin of the tumor [74]. This analysis, therefore, strengthens the proof that the sequences extracted from the genomes of these genes may be a true promoter sequence.

Table 2. Result of the CpG island assessed by sequence manipulator suite.

S/N	Gene_ID	Min. GC%	Max. GC%	Min. obs/exp	Max. obs/exp
1	APC	51.00	57.50	0.61	0.67
2	KRAS	59.50	83.00	0.78	1.14
3	TCF7L2	53.00	72.50	0.60	1.00
4	EGFR	56.00	57.00	0.62	0.90
5	IGF1R	53.00	81.50	0.61	1.22
6	CASP8	67.50	70.50	0.60	0.74
7	GNAS	67.50	70.50	0.61	0.78

Note: Min. GC%- Minimum GC content detected in each region of 1-200 base pairs; Max. GC%- Maximum GC content detected in each region of 1-200 base pairs. Min. obs/exp-Minimum observed/expected ratio of CpG dinucleotides. Max. obs/exp- Maximum observed/expected ratio of CpG dinucleotides.

3.3. Triplex Binding Interaction of the MicroRNAs and Target Genes

The specific binding properties of microRNA to proteins or (mRNA) involved in chromatin remodeling and several transcriptional regulations as broadly achieved in the literature demonstrate their multi-functionality [75]. The ability of RNA to participate in triplex formation according to Hoogsteen base pairing rules is a less studied property of microRNA [76]. In this study, the binding of the candidate microRNAs to the TSS (600 bp upstream) of the promoter region of the target genes were determined by an *in silico* analysis using the Trident software. Pasquier et al. [76] reported that formative genes are profoundly represented in the TTS, which certifies the speculation of unexpected large-scale genome regulation mediated by the triplex DNA-RNA structure. The binding of microRNAs to mRNA directly for gene expression regulation has been experimentally validated. However, research has suggested the mechanism by which microRNA forms triplex with double-stranded DNA but their exact mechanism of interaction is less well understood [26,27,77]. Also, Blanco and Montoya [78] elucidated the transient DNA/RNA protein interactions but proposed a further study of triplex formation. From the triplex-forming oligonucleotide formation result represented in Table 3, it was observed that only three of the candidate microRNAs (miR-1, miR-2, and miR-5) are capable of forming multiple triplexes along with three of the target mRNAs (KRAS, TCF7L2, and EGFR). For miR-1, four binding sites were observed in KRAS, and EGFR, while nine binding sites were determined in TCF7L2. This microRNA (miR-1) reports no hit score for the genes APC and GNAS. MicroRNAs 2 and 4 reported no triplex interaction between the target genes showing negative correlation. In microRNA-3, only one reverse (indirect) Hoogsteen pairing was observed with KRAS with hit energy of -168.93 (Figure 2). MicroRNA-5 showed two different triplex binding sites in TCF7L2 and one in both APC (indirect) and CASP8 (direct) (Figure 2). All the binding interactions in this study are of grade 5 (99 percentiles of triplex-forming interactions). Both the hit score and hit energy observed are greater than 140 and -140 respectively (Figure 2).

Table 3. Output of the results of trident showing different binding sites.

MicroRNA/Gene	KRAS	TCF7L2	APC	EGFR	CASP8	IGF1R	GNAS
miR-1	-4	-9	0	-3/+1	+1	+1	0
miR-2	0	0	0	0	0	0	0
miR-3	-1	0	0	0	0	0	0
miR-4	0	0	0	0	0	0	0
miR-5	0	-2	-1	0	+1	0	0

Legend: (0); No hit, (+); Direct Hoogsteen, (-); Reverse Hoogsteen, (Values); Number of hits; Grade: 5. Hit score: > 140; Hit energy: < -140. The heuristic score (hit score) represents Hoogsteen or Reverse Hoogsteen base pair complementarity and Thermodynamic Energy (hit energy) represents the binding energy of the triplex. The binding sites were categorized based on the number of hits relative to score and energy.

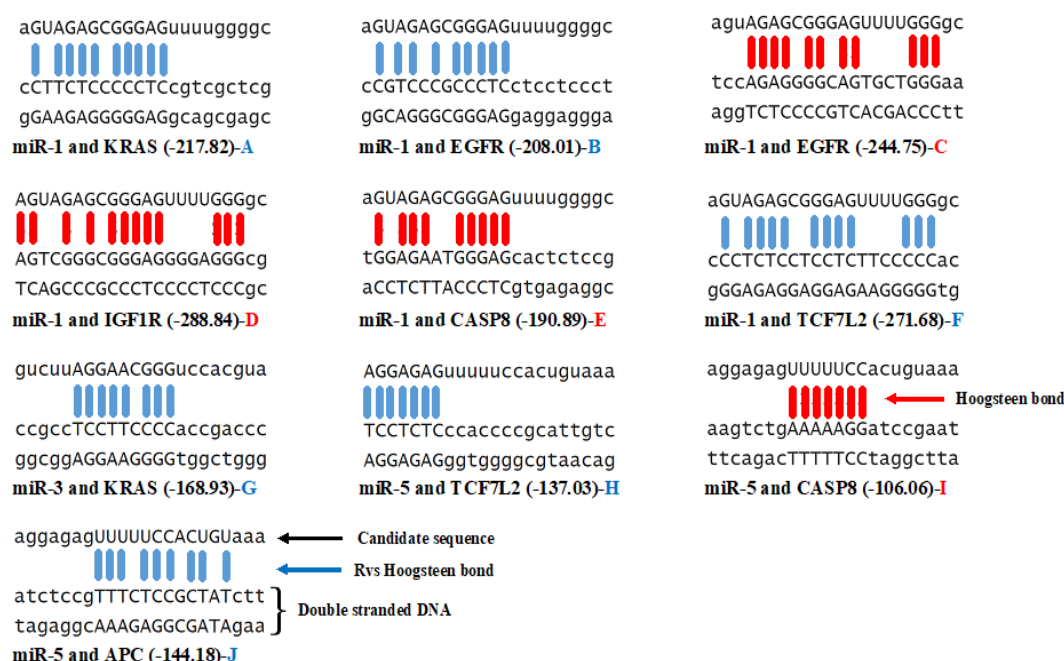


Figure 2. Structural determination of microRNA-DNA triplex formation. The first sequence in each structure (A–J) represent the candidate microRNA sequences involved in triplex binding while the two sequences without bond are the promoter sequence of the target gene. The blue bond indicates the indirect or reverse Hoogsteen bond while the red bond is the direct or Hoogsteen bond between the DNA and the microRNA. The negative values in the bracket are the hit energy of each binding.

From previous studies, it was recorded that microRNA with a 22-nucleotides sequence forms triplex structures with duplex DNA as documented by fluorescence resonance energy transfer, FRET, surface Plasmon resonance, electromobility shift assay, and Nuclear magnetic resonance [79,80]. Similar to transient protein–protein and DNA/RNA-protein interactions, Blanco and Montoya [78] suggested that transient formation of microRNA-duplex DNA triplexes may have as much biological importance as more stable interactions. Surprisingly, there are enzymes of the class helicase that possess the ability to unwind the intramolecular triplex DNA [30,81], it is feasible that this mode of action of these enzymes is similitude to which microRNA can mediate transcriptional activation [82]. On a whole, the formation of triplexes between the candidate microRNA (1, 3, and 5) and DNA double-stranded suggest that they are well conserved and crucial mechanism of transcription regulation. Comparing the thermodynamic energies of folding and triplex formation (Table 1 and Figure 2), it was observed that these microRNAs are better favoured for triplex binding to regulate gene expression in the cancer subtype than microRNA folding.

3.4. Somatic Event Evolution of the MicroRNA Target Genes

The evolutionary stage-specific and variant events of the microRNA target genes are presented in Figure 3a,b. Using the SEECancer database, all the genes involved in CRC initiation, progression, and metastasis were graphically represented and visualized using Cytoscape. From the network, four (KRAS, GNAS, APC, and EGRF) of the statistically significant genes targeted by the candidate microRNAs were indicated with yellow node. The database also reported that the other three genes (TCF7L2, CASP8, and IGF1R) are either a product of antecedent or subsequent event of the four genes implicated in CRC. TCF7L2 was seen from the database to be a subsequent event of APC mutation. Loss of heterozygosity (LOH) in the APC was shown to affect TP53 gene through mutation which in turn results in the mutation of TCF7L2. APC was found to be specific to early stage of CRC while GNAS was seen at the metastatic stage. Figure 3b shows that alteration, mutation, methylation aberrance, and loss of copy number of APC gene may result in CRC initiation, progression, or metastasis. Previous studies have reported that the loss of function, as well as methylation aberrance and mutation of the APC gene, is associated with early events in CRC [83–92].

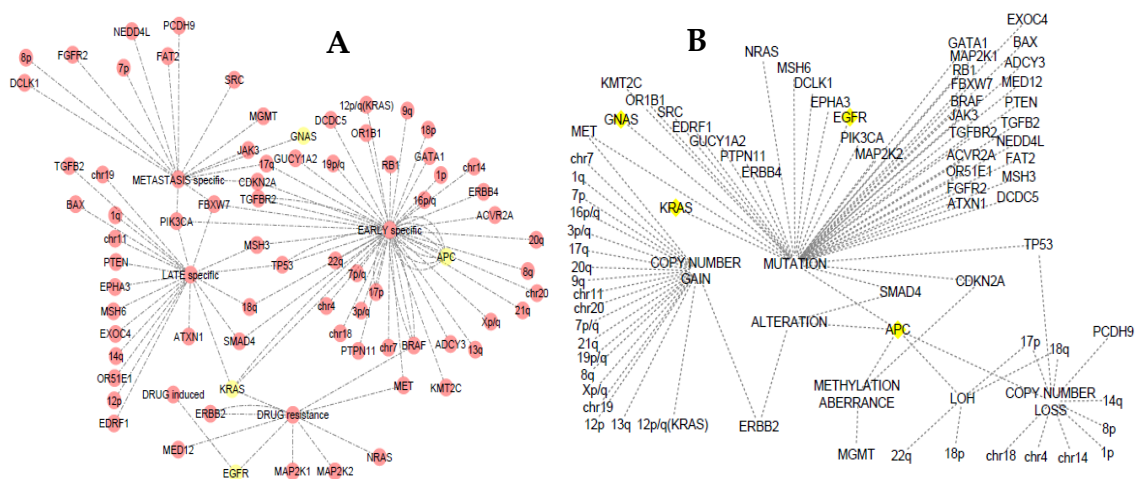


Figure 3. Evolutionary stage-specific somatic events in colorectal cancer (CRC). (A) stage-specific targets. (B) variant types and specific variants of the microRNA target genes. Yellow nodes indicate the microRNA target genes.

The antecedent event resulting in the mutation of KRAS was also detected to be from the mutation of CASP8 which in turn mutates PIK3CA gene. Mutation in the gene KRAS was revealed to be either at the early, late, or drug resistance stage. Burner et al. [93] reported the early event of KRAS in CRC as detected by DNA amplification and oligonucleotide hybridization, as well as by RNase A mismatch cleavage analysis. Using a bioinformatics approach, Youn and Simon [94] also suggested that KRAS mutations occur as the first event with high probability for both colorectal and lung tumors. Genome-editing technology also revealed that mutations in this gene are sufficient to initiate tumour progression [95]. The report of Fumagalli et al. [96] using a high throughput experiment suggest that the initiating APC and KRAS mutations drives efficient proliferation and growth, whereas inactivating mutations in SMAD4 block differentiation during tumor progression. Mutation in EGFR was shown to be at the drug-induced and resistance stage event (Figure 3a,b). In metastatic CRC using high throughput experiment, Morelli et al. [97], concluded that acquired mutations of this gene in metastatic CRC patients were correlated with acquired resistance to anti-EGFR monoclonal antibodies. Mutation in other genes targeted by the candidate microRNAs may be involved in colorectal carcinogenesis.

3.5. Co-expression Analysis

GeneMANIA was used to generate a hypothesis about the candidate microRNA target genes' significance. The co-expression analysis of the candidate microRNAs result was represented in Figure 4. The output generated genes that are functionally similar, or have shared properties with microRNA target genes, displayed an interactive functional association network, illustrating the relationships among the genes and gene of interest.

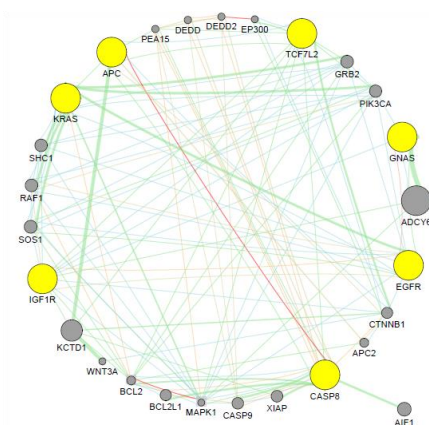


Figure 4. Co-expression analysis network of the prioritized microRNA target genes. The target genes are shown in yellow nodes while other genes in grey represent the associated genes. The red edges showed that the connected genes are co-expressed while the green edges are associated with genetic interactions. Other interactions include shared protein domains, physical interaction, pathway, and predicted interaction.

Genetic interaction of this network accounted for 43.83% of the total interactions, shared protein domain accounted for 26.28, physical interaction accounted for 14.23, co-expressed genes accounted for 10.14%. Only 3.89 and 2.62% are accounted for pathway and predicted interaction respectively. Within the microRNA target genes, only APC and CASP8 are co-expressed, although neighbouring genes found to be co-expressed are Bcl-2 and MAPK-1. Since this tool assigns a weighting feature to dataset to determine how genes in a list are well connected to each other or to determine which types of functional genomic data are most useful to retrieve for finding more genes similar to the query, KCTD1 and ADCY6 were assigned the highest weight due to size and are genetically connected with APC and GNAS respectively. Upregulation of ADCY6 activates the CREB pathway by increasing the tumorigenic potential of cells reported in gastric cancer [98]. Certain genes in the Wnt pathway influence KCTD1-mediated downregulation of β -catenin and suggested that KCTD1 functions as a novel inhibitor of Wnt signaling pathway by enhancing β -catenin degradation [99]. Inappropriate activation of this pathway has been observed in CRC [100,101]. These associated genes may then provide further a complete microRNA-gene network for CRC diagnosis and disease management.

4. Materials and Methods

The secondary structure, CpG island and triplex binding analysis, Somatic Events in Evolution of the target genes, and co-expression analysis were performed on both the candidate microRNAs and their targets using the following tools: Mfold at <http://unafold.rna.albany.edu/?q=mfold>, SMS at <http://www.bioinformatics.org/sms2/translate.html>, NCBI at <https://blast.ncbi.nlm.nih.gov/Blast.cgi>, Ensembl at <http://www.ensembl.org/index.html>, UCSC at <https://genome.ucsc.edu/>, Trident at <http://trident.stjude.org/>, SEECancer at <http://biocc.hrbmu.edu.cn/SEECancer/index.jsp>, GeneMANIA at <http://genemania.org/>, and Cytoscape at <https://cytoscape.org/>.

4.1. Datasets

Five candidate microRNAs decoded as miR-1 to 5 were identified using in silico approach and their targets genes were prioritized through three different target prediction tools (TargetScan, miRDB, and miRDIP) to generate seven genes namely: *APC*, *GNAS*, *EGFR*, *TCF7L2*, *KRAS*, *IGF1R*, and *CASP8*. The sequences of these microRNAs together with the promoter sequences of their targets were used for functional determination in this study.

4.2. Structural Determination of Candidate microRNA

The mature sequences of candidate microRNAs were submitted to Mfold program accessed at <http://www.mfold.rna.albany.edu> for predicting their secondary structure (s) and the free energy. MFold accepts a single nucleotide sequence as input mainly in FASTA format. The output file produced by MFold contains the calculated energy matrices that determine all optimal and suboptimal secondary structures for the folded nucleic acid molecule. This file is read by the companion program, PlotFold, which can display any of several different graphic representations of optimal and suboptimal secondary structures for the folded molecule.

4.3. Promoter Sequence Extraction

The database for Ensembl was accessed at <http://www.ensembl.org/index.html>. The seven prioritized microRNA target genes were used as input to generate their promoter sequences. From the display, the exons are highlighted in the pink background and red text, the sequence before exon 1 is the promoter sequence. Six-hundred base pairs 5'-flanking sequences were retrieved. To prevent disparity of the annotation concerning the transcription start site among databases, the obtained sequences were further analyzed in USCC using the BLAT tool at <https://genome-euro.ucsc.edu/>. Furthermore, CpG island feature and GC content were considered.

4.4. CpG Island Analysis

The sequence manipulator suites (SMS) report potential CpG island regions using the method described by [102]. The calculation was performed using a 200 base pairs window moving across the sequence at 1 base pairs intervals. CpG islands are defined as sequence ranges where the observed/expected value is greater than 0.6 and the GC content is greater than 50%. The expected number of CpG dimers in a window is calculated as the number of 'C's in the window multiplied by the number of 'G's in the window, divided by the window length. While the ration of the observed to the expected is calculated as $\text{Obs/Exp CpG} = \text{Number of CpG} * N / (\text{Number of C} * \text{Number of G})$, where N = length of the sequence. The CpG island suite was accessed at http://www.bioinformatics.org/sms2/cpg_islands.html for the identification of the CpG islands in the promoter sequences of the prioritized microRNA target genes.

4.5. Triplex Binding Analysis

Trident was accessed at <http://trident.stjude.org/>. This online prediction tool was used to identify triplex binding sites between the candidate microRNAs obtained in our previous study and the promoter region of the statistical significant target genes. Sequences of microRNAs and their targets were used as input. The output of this search was given based on the number of sites that the sequence of the candidate microRNA was able to bind on the promoter sequence. The grade, hit score, types of binding (direct or indirect), and hit energy were also considered.

4.6. Staging Analysis

The SEECancer database was accessed at <http://biocc.hrbmu.edu.cn/SEECancer> to explore the evolutionary stage-specific somatic events in CRC as well as their temporal order [103]. From the web interface, evolutionary stage and temporal order were selected individually to query the target genes

as well as the cancer subtype respectively. The result of these search generated a txt file as input in Cytoscape to generate a network of events based on variant types and evolutionary events in CRC.

4.7. Co-expression Analysis

A co-expression network identifies which genes have a tendency to show a coordinated expression pattern across a group of samples. This co-expression network can be represented as a gene–gene similarity matrix, which can be used in downstream analyses [104]. Seven prioritized genes target by candidate microRNA were used as input to generate the co-expressed genes using GeneMANIA <http://www.genemania.org> (<http://apps.cytoscape.org/apps/genemania>) plugin of Cytoscape. This database is used to generate hypotheses about gene function, analyzing gene lists and prioritizing genes for functional assays. The output is a network of genes with several interactions.

4.8. Statistical Analysis

For microRNA selection, the BlastN parameters were set at 1e-2 for expected value, 7.0 for word size, and 90–99% for similarity index. The CD-HIT-EST-2D parameters were set at 0.90 for threshold and 7.0 for word size. The genes considered in DAVID v6.8 were regarded statistically significant at p -value of 1.8×10^3 with the Benjamini score of 1.6×10^{-2} . A value of $p < 0.05$ was considered to indicate a statistically significant difference.

4.9. Data Availability

The datasets and the clinical data were obtained from the online databases as described above in the methods and their websites are as follows: <http://www.mirbase.org/> for reference dataset, microRNA associated with CRC at cancer were <http://www.picb.ac.cn/dbDEM/>, <http://www.mir2disease.org/>, <http://www.cuilab.cn/hmdd>, and <http://mircancer.ecu.edu/>. While Mfold, SMS, NCBI, Ensembl, UCSC, Trident, SEECancer, GeneMANIA and Cytoscape were accessed from the server.

5. Conclusions

Understanding microRNA's secondary structures, thermodynamic parameters, and targets may deliver greater promise towards their diagnostic potentials and mechanisms in the management of CRC. The secondary structure together with the thermodynamic parameters of the candidate microRNAs may, therefore, provide a valid result regarding their target even when the conservation of the microRNA is unknown. Also, the secondary structure of the candidate microRNAs suggests a conformational role to modulate target interactions and therefore can be used to explain the different degree of genetic regulation in CRC.

MicroRNAs that are capable of triplex formation with duplex DNA are more frequently positively correlated with gene transcripts. MiR-1, miR-3, and miR-5 are suggested to have significant enrichment of positive correlation with the target gene involved in the triplex structure. This analysis further confirmed that the targets of the candidate microRNAs are enriched with microRNA triplex binding sites. Furthermore, microRNA function may then depend not only on their sequences but also their structures and triplex binding interaction with their targets. Hypothetically speaking, microRNAs targeting these genes can be inferred to be an important regulator in the stage-specific events. Since miR-1, miR-2, and miR-3 are regulators of APC, it can be concluded that these microRNAs may be used as a diagnostic tool in the early detection of CRC. Also, miR-1 regulating GNAS may be involved in metastatic stage-specific event in CRC. MicroRNAs regulating KRAS may also be involved in early, late, or drug-resistant stage-specific events in CRC (MiR-1,3 and 5). However, further molecular experiments are on-going to confirm the function of these identified microRNAs alongside their targets in CRC.

Author Contributions: All authors have made significant contributions to the submission of the article. A.O.F. conceived the concept and the design of the manuscript, A.K. and A.P. provided the necessary software required

to carry out the analysis. The analysis and data interpretations were done by both A.O.F. and A.P. while A.K. drafted the rough draft and also substantively revised the manuscript. Finally, all authors (A.O.F., A.K., and A.P.) read and approved the submitted version of the manuscript for publication. Also, A.O.F., A.K., and A.P. agreed to be personally accountable for their personal contributions and ensured that questions related to the accuracy or integrity of any part of the work, even ones in which the author was not personally involved, are appropriately investigated, resolved, and the resolution documented in the literature.

Funding: This research received no external funding.

Acknowledgments: Authors will like to acknowledge the Plant Omics Laboratory and the Bioinformatics Research Group of the University of the Western Cape for their usual support. Also, Taiwo Akinsoji (MBBS, M.PH) University of Illinois at Springfield for co-proof reading the final manuscript.

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

Abbreviations

miRNA	microRNA
CRC	Colorectal cancer
MRNA	Target genes
δG	Free energy in plot profile
ΔG	Optimal energy of the secondary structure

References

1. Kuo, T.Y.; Hsi, E.; Yang, I.P.; Tsai, P.C.; Wang, J.Y.; Juo, S.H.H. Computational analysis of mRNA expression profiles identifies microRNA-29a/c as predictor of colorectal cancer early recurrence. *PLoS ONE* **2012**, *7*, e31587. [[CrossRef](#)] [[PubMed](#)]
2. Wolpin, B.M.; Meyerhardt, J.A.; Mamon, H.J.; Mayer, R.J. Adjuvant treatment of colorectal cancer. *Cancer J. Clin.* **2007**, *57*, 1681–1685. [[CrossRef](#)] [[PubMed](#)]
3. Longo, W.E.; Johnson, F.E. The preoperative assessment and postoperative surveillance of patients with colon and rectal cancer. *Surg. Clin. North Am.* **2002**, *82*, 1091–1108. [[CrossRef](#)]
4. Fadaka, A.O.; Pretorius, A.; Klein, A. Biomarkers for Stratification in Colorectal Cancer: MicroRNAs. *Cancer Control* **2019**, *26*. [[CrossRef](#)] [[PubMed](#)]
5. Schickel, R.; Boyerinas, B.; Park, S.; Peter, M. MicroRNAs: Key players in the immune system, differentiation, tumorigenesis and cell death. *Oncogene* **2008**, *27*, 5959. [[CrossRef](#)]
6. Hammond, S.M. MicroRNAs as tumor suppressors. *Nat. Genet.* **2007**, *39*, 582. [[CrossRef](#)]
7. Liu, W.; Mao, S.Y.; Zhu, W.Y. Impact of tiny miRNAs on cancers. *World J. Gastroenterol.* **2007**, *13*, 497. [[CrossRef](#)]
8. Iorio, M.V.; Croce, C.M. MicroRNAs in cancer: Small molecules with a huge impact. *J. Clin. Oncol.* **2009**, *27*, 5848. [[CrossRef](#)]
9. Bartel, D.P. MicroRNAs: Genomics, biogenesis, mechanism, and function. *Cell* **2004**, *116*, 2812–2897. [[CrossRef](#)]
10. Brennecke, J.; Hipfner, D.R.; Stark, A.; Russell, R.B.; Cohen, S.M. Bantam encodes a developmentally regulated microRNA that controls cell proliferation and regulates the proapoptotic gene hid in Drosophila. *Cell* **2003**, *113*, 253–256. [[CrossRef](#)]
11. Chen, C.Z.; Li, L.; Lodish, H.F.; Bartel, D.P. MicroRNAs modulate hematopoietic lineage differentiation. *Science* **2004**, *303*, 83–86. [[CrossRef](#)] [[PubMed](#)]
12. Fadaka, A.O.; Ojo, B.A.; Adewale, O.B.; Esho, T.; Pretorius, A. Effect of dietary components on miRNA and colorectal carcinogenesis. *Cancer Cell Int.* **2018**, *18*, 130. [[CrossRef](#)]
13. Lee, Y.S.; Dutta, A. MicroRNAs in cancer. *Annu. Rev. Pathol. Mech. Dis.* **2009**, *4*, 199–227. [[CrossRef](#)] [[PubMed](#)]
14. Motoyama, K.; Inoue, H.; Takatsuno, Y.; Tanaka, F.; Mimori, K.; Uetake, H.; Sugihara, K.; Mori, M. Over- and under-expressed microRNAs in human colorectal cancer. *Int. J. Oncol.* **2009**, *34*, 1069–1075. [[PubMed](#)]
15. Aslam, M.; Taylor, K.; Pringle, J.; Jameson, J. MicroRNAs are novel biomarkers of colorectal cancer. *Br. J. Surg.* **2009**, *96*, 702–710. [[CrossRef](#)] [[PubMed](#)]
16. Manne, U.; Shanmugam, C.; Bovell, L.; Katkooori, V.R.; Bumpers, H.L. miRNAs as biomarkers for management of patients with colorectal cancer. *Biomark. Med.* **2010**, *4*, 7617–7670. [[CrossRef](#)] [[PubMed](#)]

17. Xi, Y.; Edwards, J.R.; Ju, J. Investigation of miRNA biology by bioinformatic tools and impact of miRNAs in colorectal cancer—Regulatory relationship of c-Myc and p53 with miRNAs. *Cancer Inform.* **2007**, *3*. [[CrossRef](#)]
18. Ohler, U.; Niemann, H. Identification and analysis of eukaryotic promoters: Recent computational approaches. *Trends Genet.* **2001**, *17*, 56–60. [[CrossRef](#)]
19. Moser, H.E.; Dervan, P.B. Sequence-specific cleavage of double helical DNA by triple helix formation. *Science* **1987**, *238*, 645–650. [[CrossRef](#)]
20. Le Doan, T.; Perrouault, L.; Praseuth, D.; Habhoub, N.; Decout, J.L.; Thuong, N.T.; Lhomme, J.; Héène, C. Sequence-specific recognition, photocrosslinking and cleavage of the DNA double helix by an oligo-(α)-thymidylate covalently linked to an azidoproflavine derivative. *Nucleic Acids Res.* **1987**, *15*, 77497–77760. [[CrossRef](#)]
21. Broitman, S.L.; Im, D.D.; Fresco, J.R. Formation of the triple-stranded polynucleotide helix, poly (AAU). *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 5120–5124. [[CrossRef](#)] [[PubMed](#)]
22. Wells, R.; Collier, D.; Hanvey, J.; Shimizu, M.; Wohlrab, F. The chemistry and biology of unusual DNA structures adopted by oligopurine. oligopyrimidine sequences. *FASEB J.* **1988**, *2*, 29392–29949. [[CrossRef](#)]
23. Buske, F.A.; Mattick, J.S.; Bailey, T.L. Potential in vivo roles of nucleic acid triple-helices. *RNA Biol.* **2011**, *8*, 427–439. [[CrossRef](#)] [[PubMed](#)]
24. Jalali, S.; Singh, A.; Maiti, S.; Scaria, V. Genome-wide computational analysis of potential long noncoding RNA mediated DNA: RNA triplexes in the human genome. *J. Transl. Med.* **2017**, *15*, 186. [[CrossRef](#)]
25. Li, Y.; Syed, J.; Sugiyama, H. RNA-DNA triplex formation by long noncoding RNAs. *Cell Chem. Biol.* **2016**, *23*, 13251–13333. [[CrossRef](#)] [[PubMed](#)]
26. Janowski, B.A.; Younger, S.T.; Hardy, D.B.; Ram, R.; Huffman, K.E.; Corey, D.R. Activating gene expression in mammalian cells with promoter-targeted duplex RNAs. *Nat. Chem. Biol.* **2007**, *3*, 166. [[CrossRef](#)]
27. Check, E. RNA interference: Hitting the on switch. *Nature* **2007**, *448*, 855–858. [[CrossRef](#)]
28. Kanak, M.; Alseiyari, M.; Balasubramanian, P.; Addanki, K.; Aggarwal, M.; Noorali, S.; Kalsum, A.; Mahalingam, K.; Pace, G.; Panasik, N. Triplex-forming MicroRNAs form stable complexes with HIV-1 provirus and inhibit its replication. *Appl. Immunohistochem. Mol. Morphol.* **2010**, *18*, 5325–5345. [[CrossRef](#)]
29. Schmitz, K.-M.; Mayer, C.; Postepska, A.; Grummt, I. Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes Dev.* **2010**, *24*, 2264–2269. [[CrossRef](#)]
30. Goñi, J.R.; De La Cruz, X.; Orozco, M. Triplex-forming oligonucleotide target sequences in the human genome. *Nucleic Acids Res.* **2004**, *32*, 3543–3560. [[CrossRef](#)]
31. Jenjaroenpun, P.; Kuznetsov, V.A. In TTS mapping: Integrative WEB tool for analysis of triplex formation target DNA sequences, G-quadruplets and non-protein coding regulatory DNA elements in the human genome. *BMC Genom.* **2009**, *S9*. [[CrossRef](#)] [[PubMed](#)]
32. Jenjaroenpun, P.; Chew, C.S.; Yong, T.P.; Choowongkamon, K.; Thammasorn, W.; Kuznetsov, V.A. The TTSMI database: A catalog of triplex target DNA sites associated with genes and regulatory elements in the human genome. *Nucleic Acids Res.* **2014**, *43*, D110–D116. [[CrossRef](#)] [[PubMed](#)]
33. Buske, F.A.; Bauer, D.C.; Mattick, J.S.; Bailey, T.L. Triplexator: Detecting nucleic acid triple helices in genomic and transcriptomic data. *Genome Res.* **2012**, *22*, 1372–1381. [[CrossRef](#)] [[PubMed](#)]
34. Wu, Q.; Gaddis, S.S.; MacLeod, M.C.; Walborg, E.F.; Thames, H.D.; DiGiovanni, J.; Vasquez, K.M. High-affinity triplex-forming oligonucleotide target sequences in mammalian genomes. *Mol. Carcinog* **2007**, *46*, 152–153. [[CrossRef](#)] [[PubMed](#)]
35. Battle, E.; Bacani, J.; Begthel, H.; Jonkeer, S.; Gregorieff, A.; van de Born, M.; Malats, N.; Sancho, E.; Boon, E.; Pawson, T. EphB receptor activity suppresses colorectal cancer progression. *Nature* **2005**, *435*, 1126. [[CrossRef](#)] [[PubMed](#)]
36. Dow, L.E.; O'Rourke, K.P.; Simon, J.; Tschaharganeh, D.F.; van Es, J.H.; Clevers, H.; Lowe, S.W. Apc restoration promotes cellular differentiation and reestablishes crypt homeostasis in colorectal cancer. *Cell* **2015**, *161*, 15391–15552. [[CrossRef](#)] [[PubMed](#)]
37. Scholer-Dahirel, A.; Schlabach, M.R.; Loo, A.; Bagdasarian, L.; Meyer, R.; Guo, R.; Woolfenden, S.; Kristine, K.Y.; Markovits, J.; Killary, K. Maintenance of adenomatous polyposis coli (APC)-mutant colorectal cancer is dependent on Wnt/ β -catenin signaling. *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 17135–17140. [[CrossRef](#)]

38. Misale, S.; Yaeger, R.; Hobor, S.; Scala, E.; Janakiraman, M.; Liska, D.; Valtorta, E.; Schiavo, R.; Buscarino, M.; Siravegna, G. Emergence of KRAS mutations and acquired resistance to anti-EGFR therapy in colorectal cancer. *Nature* **2012**, *486*, 532. [[CrossRef](#)]
39. De Roock, W.; De Vriendt, V.; Normanno, N.; Ciardiello, F.; Tejpar, S. KRAS, BRAF, PIK3CA, and PTEN mutations: Implications for targeted therapies in metastatic colorectal cancer. *Lancet Oncol.* **2011**, *12*, 594–603. [[CrossRef](#)]
40. Bass, A.J.; Lawrence, M.S.; Brace, L.E.; Ramos, A.H.; Drier, Y.; Cibulskis, K.; Sougnez, C.; Voet, D.; Saksena, G.; Sivachenko, A. Genomic sequencing of colorectal adenocarcinomas identifies a recurrent VTI1A-TCF7L2 fusion. *Nat. Genet.* **2011**, *43*, 964. [[CrossRef](#)]
41. Theodoropoulos, G.E.; Gazouli, M.; Vaiopoulou, A.; Leandrou, M.; Nikouli, S.; Vassou, E.; Kouraklis, G.; Nikiteas, N. Polymorphisms of caspase 8 and caspase 9 gene and colorectal cancer susceptibility and prognosis. *Int. J. Colorectal Dis.* **2011**, *26*, 1113–1118. [[CrossRef](#)] [[PubMed](#)]
42. Inno, A.; Di Salvatore, M.; Cenci, T.; Martini, M.; Orlandi, A.; Strippoli, A.; Ferrara, A.M.; Bagalà, C.; Cassano, A.; Larocca, L.M. Is there a role for IGF1R and c-MET pathways in resistance to cetuximab in metastatic colorectal cancer? *Clin. Colorectal Cancer* **2011**, *10*, 325–332. [[CrossRef](#)] [[PubMed](#)]
43. Idziaszczyk, S.; Wilson, C.; Smith, C.; Adams, D.; Cheadle, J. Analysis of the frequency of GNAS codon 201 mutations in advanced colorectal cancer. *Cancer Genet. Cytogenetics* **2010**, *202*, 676–679. [[CrossRef](#)] [[PubMed](#)]
44. Kojima, M.; Sudo, H.; Kawachi, J.; Takizawa, S.; Kondou, S.; Nobumasa, H.; Ochiai, A. MicroRNA markers for the diagnosis of pancreatic and biliary-tract cancers. *PLoS ONE* **2015**, *10*, e0118220. [[CrossRef](#)]
45. Plieskatt, J.L.; Rinaldi, G.; Feng, Y.; Peng, J.; Yonglitthipagon, P.; Easley, S.; Laha, T.; Pairojkul, C.; Bhudhisawasdi, V.; Sripa, B. Distinct miRNA signatures associate with subtypes of cholangiocarcinoma from infection with the tumorigenic liver fluke *Opisthorchis viverrini*. *J. Hepatol.* **2014**, *61*, 8508–8558. [[CrossRef](#)]
46. Taylor, B.S.; Schultz, N.; Hieronymus, H.; Gopalan, A.; Xiao, Y.; Carver, B.S.; Arora, V.K.; Kaushik, P.; Cerami, E.; Reva, B. Integrative genomic profiling of human prostate cancer. *Cancer Cell* **2010**, *18*, 11–22. [[CrossRef](#)]
47. Ng, S.-B.; Yan, J.; Huang, G.; Selvarajan, V.; Tay, J.L.-S.; Lin, B.; Bi, C.; Tan, J.; Kwong, Y.-L.; Shimizu, N. Dysregulated microRNAs affect pathways and targets of biologic relevance in nasal-type natural killer/T-cell lymphoma. *Blood* **2011**, *118*, 49194–49929. [[CrossRef](#)]
48. Zhang, P.; Bill, K.; Liu, J.; Young, E.; Peng, T.; Bolshakov, S.; Hoffman, A.; Song, Y.; Demicco, E.G.; Terrada, D.L. MiR-155 is a liposarcoma oncogene that targets casein kinase-1 α and enhances β -catenin signaling. *Cancer Res.* **2012**, *72*, 17511–17762. [[CrossRef](#)]
49. Munding, J.B.; Adai, A.T.; Maghnouj, A.; Urbanik, A.; Zöllner, H.; Liffers, S.T.; Chromik, A.M.; Uhl, W.; Szafranska-Schwarzbach, A.E.; Tannapfel, A. Global microRNA expression profiling of microdissected tissues identifies miR-135b as a novel biomarker for pancreatic ductal adenocarcinoma. *Int. J. Cancer* **2012**, *131*, E86–E95. [[CrossRef](#)]
50. Zhang, N.; Wang, X.; Huo, Q.; Li, X.; Wang, H.; Schneider, P.; Hu, G.; Yang, Q. The oncogene metadherin modulates the apoptotic pathway based on the tumor necrosis factor superfamily member TRAIL (Tumor Necrosis Factor-related Apoptosis-inducing Ligand) in breast cancer. *J. Biol. Chem.* **2013**, *288*, 9396–9407. [[CrossRef](#)]
51. Biagioni, F.; Ben-Moshe, N.B.; Fontemaggi, G.; Canu, V.; Mori, F.; Antoniani, B.; Di Benedetto, A.; Santoro, R.; Germoni, S.; De Angelis, F. miR-10b*, a master inhibitor of the cell cycle, is down-regulated in human breast tumours. *EMBO Mol. Med.* **2012**, *4*, 12141–12229. [[CrossRef](#)] [[PubMed](#)]
52. Özata, D.M.; Caramuta, S.; Velázquez-Fernández, D.; Akçakaya, P.; Xie, H.; Höög, A.; Zedenius, J.; Bäckdahl, M.; Larsson, C.; Lui, W.-O. The role of microRNA deregulation in the pathogenesis of adrenocortical carcinoma. *Endocr. Relat. Cancer* **2011**, *18*, 6436–6455. [[CrossRef](#)] [[PubMed](#)]
53. Namløs, H.M.; Meza-Zepeda, L.A.; Barøy, T.; Østensen, I.H.; Kresse, S.H.; Kuijjer, M.L.; Serra, M.; Bürger, H.; Cleton-Jansen, A.-M.; Myklebost, O. Modulation of the osteosarcoma expression phenotype by microRNAs. *PLoS ONE* **2012**, *7*, e48086. [[CrossRef](#)] [[PubMed](#)]
54. Jung, H.M.; Phillips, B.L.; Patel, R.S.; Cohen, D.M.; Jakymiw, A.; Kong, W.W.; Cheng, J.Q.; Chan, E.K. Keratinization-associated miR-7 and miR-21 regulate tumor suppressor reversion-inducing cysteine-rich protein with kazal motifs (RECK) in oral cancer. *J. Biol. Chem.* **2012**, *287*, 29261–29272. [[CrossRef](#)]

55. Kretschmer-Kazemi Far, R.; Sczakiel, G. The activity of siRNA in mammalian cells is related to structural target accessibility: A comparison with antisense oligonucleotides. *Nucleic Acids Res.* **2003**, *31*, 44174–44424. [[CrossRef](#)]
56. Overhoff, M.; Alken, M.; Far, R.K.-K.; Lemaitre, M.; Lebleu, B.; Sczakiel, G.; Robbins, I. Local RNA target structure influences siRNA efficacy: A systematic global analysis. *J. Mol. Biol.* **2005**, *348*, 8718–8781. [[CrossRef](#)]
57. Mückstein, U.; Tafer, H.; Hackermüller, J.; Bernhart, S.H.; Stadler, P.F.; Hofacker, I.L. Thermodynamics of RNA–RNA binding. *Bioinform.* **2006**, *22*, 1177–1182. [[CrossRef](#)]
58. Belter, A.; Gudanis, D.; Rolle, K.; Piwecka, M.; Gdaniec, Z.; Naskręć-Barciszewska, M.Z.; Barciszewski, J. Mature miRNAs form secondary structure, which suggests their function beyond RISC. *PLoS ONE* **2014**, *9*, e113848. [[CrossRef](#)]
59. Maiti, M.; Nauwelaerts, K.; Lescrinier, E.; Schuit, F.C.; Herdewijn, P. Self-complementary sequence context in mature miRNAs. *Biochem. Biophys. Res. Commun.* **2010**, *392*, 5725–5776. [[CrossRef](#)]
60. Belter, A.; Naskręć-Barciszewska, M. The correlation of structural features of mature miRNAs with their biological function. *Biotechnol. J. Biotechnol. Comput. Biol. Bionanotechnol.* **2014**, *95*, 187–191.
61. Ronchieri, E.; D’Agostino, D.; Milanesi, L.; Merelli, I. microRNA–target interaction: A parallel approach for computing pairing energy. In *Proceedings of the 2016 24th Euromicro International Conference on Parallel, Distributed, and Network-Based Processing (PDP), Heraklion, Greece, 17–19 February 2016*; IEEE Press: Piscataway, NJ, USA, 2016; pp. 535–540.
62. Witkos, T.M.; Koscińska, E.; Krzyżosiak, W.J. Practical aspects of microRNA target prediction. *Curr. Mol. Med.* **2011**, *11*, 93–109. [[CrossRef](#)] [[PubMed](#)]
63. Betel, D.; Koppal, A.; Agius, P.; Sander, C.; Leslie, C. Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. *Genome Biol.* **2010**, *11*, R90. [[CrossRef](#)]
64. Kertesz, M.; Iovino, N.; Unnerstall, U.; Gaul, U.; Segal, E. The role of site accessibility in microRNA target recognition. *Nat. Genet.* **2007**, *39*, 1278. [[CrossRef](#)]
65. Arif, M.; Ochoa-Corona, F. Comparative assessment of 5′ A/T-rich overhang sequences with optimal and sub-optimal primers to increase PCR yields and sensitivity. *Mol. Biotechnol.* **2013**, *55*, 172–176. [[CrossRef](#)] [[PubMed](#)]
66. Esteller, M. Epigenetics in cancer. *New Engl. J. Med.* **2008**, *358*, 1148–1159. [[CrossRef](#)] [[PubMed](#)]
67. Oberley, M.J.; Farnham, P.J. Probing chromatin immunoprecipitates with CpG-island microarrays to identify genomic sites occupied by DNA-binding proteins. *Methods Enzym.* **2003**, *371*, 577–596.
68. Curtin, K.; Slaterry, M.L.; Samowitz, W.S. CpG island methylation in colorectal cancer: Past, present and future. *Pathol. Res. Int.* **2011**. [[CrossRef](#)] [[PubMed](#)]
69. Bird, A.P. Gene expression: DNA methylation—How important in gene control? *Nature* **1984**, *307*, 503. [[CrossRef](#)]
70. Bird, A.P. DNA methylation versus gene expression. *Development* **1984**, *83*, 31–40.
71. Takai, D.; Jones, P.A. Comprehensive analysis of CpG islands in human chromosomes 21 and 22. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 37403–37745. [[CrossRef](#)]
72. Gatter, K. Handbook of immunohistochemistry and in situ hybridisation of human carcinomas: Molecular genetics—lung and breast carcinomas. *Br. J. Cancer* **2005**, *93*, 1318. [[CrossRef](#)]
73. Portela, A.; Esteller, M. Epigenetic modifications and human disease. *Nat. Biotechnol.* **2010**, *28*, 1057. [[CrossRef](#)] [[PubMed](#)]
74. Esteller, M. CpG island hypermethylation and tumor suppressor genes: A booming present, a brighter future. *Oncogene* **2002**, *21*, 5427. [[CrossRef](#)] [[PubMed](#)]
75. Elliott, D.; Lodomery, M. *Molecular Biology of RNA*; Oxford University Press: Oxford, UK, 2017.
76. Pasquier, C.; Agnel, S.; Robichon, A. The Mapping of Predicted Triplex DNA: RNA in the Drosophila Genome Reveals a Prominent Location in Development-and Morphogenesis-Related Genes. *Genes Genomes Genet.* **2017**, *7*, 2295–2304.
77. Park, C.W.; Zeng, Y.; Zhang, X.; Subramanian, S.; Steer, C.J. Mature microRNAs identified in highly purified nuclei from HCT116 colon cancer cells. *RNA Biol.* **2010**, *7*, 606–614. [[CrossRef](#)] [[PubMed](#)]
78. Blanco, F.J.; Montoya, G. Transient DNA/RNA–protein interactions. *FEBS J.* **2011**, *278*, 16431–16650. [[CrossRef](#)] [[PubMed](#)]

79. Semerad, C.L.; James Maher III, L. Exclusion of RNA strands from a purine motif triple helix. *Nucleic Acids Res.* **1994**, *22*, 53215–53325. [[CrossRef](#)]
80. Roberts, R.W.; Crothers, D.M. Stability and properties of double and triple helices: Dramatic effects of RNA or DNA backbone composition. *Science* **1992**, *258*, 1463–1466. [[CrossRef](#)]
81. Maine, I.P.; Kodadek, T. Efficient unwinding of triplex DNA by a DNA helicase. *Biochem. Biophys. Res. Commun.* **1994**, *204*, 1119–1124. [[CrossRef](#)]
82. Paugh, S.W.; Coss, D.R.; Bao, J.; Laudermilk, L.T.; Grace, C.R.; Ferreira, A.M.; Waddell, M.B.; Ridout, G.; Naeve, D.; Leuze, M. MicroRNAs form triplexes with double stranded DNA at sequence-specific binding sites; a eukaryotic mechanism via which microRNAs could directly alter gene expression. *PLoS Comput. Biol.* **2016**, *12*, e1004744. [[CrossRef](#)]
83. Umetani, N.; Sasaki, S.; Watanabe, T.; Shinozaki, M.; Matsuda, K.; Ishigami, H.; Ueda, E.; Muto, T. Genetic alterations in ulcerative colitis-associated neoplasia focusing on APC, K-ras gene and microsatellite instability. *Jpn. J. Cancer Res.* **1999**, *90*, 1081–1087. [[CrossRef](#)] [[PubMed](#)]
84. Matkowskyj, K.A.; Chen, Z.E.; Rao, M.S.; Yang, G.-Y. Dysplastic lesions in inflammatory bowel disease: Molecular pathogenesis to morphology. *Arch. Pathol. Lab. Med.* **2013**, *137*, 338–350. [[CrossRef](#)] [[PubMed](#)]
85. Redston, M.S.; Papadopoulos, N.; Caldas, C.; Kinzler, K.W.; Kern, S.E. Common occurrence of APC and K-ras gene mutations in the spectrum of colitis-associated neoplasias. *Gastroenterology* **1995**, *108*, 3833–3892. [[CrossRef](#)]
86. Yashiro, M. Molecular alterations of colorectal cancer with inflammatory bowel disease. *Dig. Dis. Sci.* **2015**, *60*, 2251–2263. [[CrossRef](#)]
87. Molinaro, A.; Terdiman, J.P.; Willenbacher, R.; Chang, F.; Baretton, G.B.; Waldman, F.; Aust, A.E.; Loehrs, U. The APC/beta-catenin pathway in ulcerative colitis-related colorectal carcinomas: A mutational analysis. *Lancet* **2002**. [[CrossRef](#)]
88. Segditsas, S.; Rowan, A.; Howarth, K.; Jones, A.; Leedham, S.; Wright, N.; Gorman, P.; Chambers, W.; Domingo, E.; Roylance, R. APC and the three-hit hypothesis. *Oncogene* **2009**, *28*, 146. [[CrossRef](#)]
89. Kim, T.-M.; An, C.H.; Rhee, J.-K.; Jung, S.-H.; Lee, S.H.; Baek, I.-P.; Kim, M.S.; Lee, S.H.; Chung, Y.-J. Clonal origins and parallel evolution of regionally synchronous colorectal adenoma and carcinoma. *Oncotarget* **2015**, *6*, 27725. [[CrossRef](#)]
90. Sakai, E.; Fukuyo, M.; Matsusaka, K.; Ohata, K.; Doi, N.; Takane, K.; Matsushashi, N.; Fukushima, J.; Nakajima, A.; Kaneda, A. TP 53 mutation at early stage of colorectal cancer progression from two types of laterally spreading tumors. *Cancer Sci.* **2016**, *107*, 8208–8227. [[CrossRef](#)]
91. Fodde, R.; Kuipers, J.; Rosenberg, C.; Smits, R.; Kielman, M.; Gaspar, C.; van Es, J.H.; Breukel, C.; Wiegant, J.; Giles, R.H.; et al. Mutations in the APC tumour suppressor gene cause chromosomal instability. *Nat. Cell Biol.* **2001**, *3*, 433–438. [[CrossRef](#)]
92. Smolle, M.A.; Pichler, M.; Haybaeck, J.; Gerger, A. Genetic markers of recurrence in colorectal cancer. *Pharmacogenomics* **2015**, *16*, 1315–1328. [[CrossRef](#)]
93. Burmer, G.C.; Rabinovitch, P.S.; Loeb, L.A. Analysis of c-Ki-ras mutations in human colon carcinoma by cell sorting, polymerase chain reaction, and DNA sequencing. *Cancer Res.* **1989**, *49*, 2141–2146. [[PubMed](#)]
94. Youn, A.; Simon, R. Estimating the order of mutations during tumorigenesis from tumor genome sequencing data. *Bioinformatics* **2012**, *28*, 15551–15561. [[CrossRef](#)]
95. Dienstmann, R.; Vermeulen, L.; Guinney, J.; Kopetz, S.; Tejpar, S.; Tabernero, J. Consensus molecular subtypes and the evolution of precision medicine in colorectal cancer. *Nat. Rev. Cancer* **2017**, *17*, 79. [[CrossRef](#)] [[PubMed](#)]
96. Fumagalli, A.; Drost, J.; Suijkerbuijk, S.J.; Van Boxtel, R.; De Ligt, J.; Offerhaus, G.J.; Begthel, H.; Beerling, E.; Tan, E.H.; Sansom, O.J. Genetic dissection of colorectal cancer progression by orthotopic transplantation of engineered cancer organoids. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, E2357–E2364. [[CrossRef](#)] [[PubMed](#)]
97. Morelli, M.; Overman, M.; Dasari, A.; Kazmi, S.; Mazard, T.; Vilar, E.; Morris, V.; Lee, M.; Herron, D.; Eng, C. Characterizing the patterns of clonal selection in circulating tumor DNA from patients with colorectal cancer refractory to anti-EGFR treatment. *Ann. Oncol.* **2015**, *26*, 7317–7336. [[CrossRef](#)]
98. Hong, S.H.; Goh, S.H.; Lee, S.J.; Hwang, J.A.; Lee, J.; Choi, I.J.; Seo, H.; Park, J.-H.; Suzuki, H.; Yamamoto, E. Upregulation of adenylate cyclase 3 (ADCY3) increases the tumorigenic potential of cells by activating the CREB pathway. *Oncotarget* **2013**, *4*, 1791. [[CrossRef](#)] [[PubMed](#)]

99. Li, X.; Chen, C.; Wang, F.; Huang, W.; Liang, Z.; Xiao, Y.; Wei, K.; Wan, Z.; Hu, X.; Xiang, S. KCTD1 suppresses canonical Wnt signaling pathway by enhancing β -catenin degradation. *PLoS ONE* **2014**, *9*, e94343. [[CrossRef](#)] [[PubMed](#)]
100. Cui, J.; Zhou, X.; Liu, Y.; Tang, Z.; Romeih, M. Wnt signaling in hepatocellular carcinoma: Analysis of mutation and expression of beta-catenin, T-cell factor-4 and glycogen synthase kinase 3-beta genes. *J. Gastroenterol. Hepatol.* **2003**, *18*, 2802–2887. [[CrossRef](#)]
101. Ewan, K.B.; Dale, T.C. The potential for targeting oncogenic wnt/ β -catenin signaling in therapy. *Curr. Drug Targets* **2008**, *9*, 5325–5347. [[CrossRef](#)]
102. Gardiner-Garden, M.; Frommer, M. CpG islands in vertebrate genomes. *J. Mol. Biol.* **1987**, *196*, 2612–2682. [[CrossRef](#)]
103. Zhang, H.; Luo, S.; Zhang, X.; Liao, J.; Quan, F.; Zhao, E.; Zhou, C.; Yu, F.; Yin, W.; Zhang, Y. SEECancer: A resource for somatic events in evolution of cancer genome. *Nucleic Acids Res.* **2017**, *46*, 1018–1026. [[CrossRef](#)] [[PubMed](#)]
104. Van Dam, S.; Vosa, U.; van der Graaf, A.; Franke, L.; de Magalhaes, J.P. Gene co-expression analysis for functional classification and gene–disease predictions. *Brief. Bioinform.* **2017**, *19*, 5755–5792. [[CrossRef](#)] [[PubMed](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).