Published in final edited form as:
Adv Drug Deliv Rev. 2015 January ; 0: 94–103. doi:10.1016/j.addr.2014.10.030.

Contribution of Bioinformatics prediction in microRNA-based cancer therapeutics

Jasjit K Banwait[a] and **Dhundy R Bastola**[a,*]

[a]College of Information Science and Technology, University of Nebraska at Omaha, 1110 South 67th Street, PKI 172, Omaha, Nebraska, 68106, USA

Abstract

Despite enormous efforts, cancer remains one of the most lethal diseases in the world. With the advancement of high throughput technologies massive amounts of cancer data can be accessed and analyzed. Bioinformatics provides a platform to assist biologists in developing minimally invasive biomarkers to detect cancer, and in designing effective personalized therapies to treat cancer patients. Still, the early diagnosis, prognosis, and treatment of cancer are an open challenge for the research community. MicroRNAs (miRNAs) are small non-coding RNAs that serve to regulate gene expression. The discovery of deregulated miRNAs in cancer cells and tissues has led many to investigate the use of miRNAs as potential biomarkers for early detection, and as a therapeutic agent to treat cancer. Here we describe advancements in computational approaches to predict miRNAs and their targets, and discuss the role of bioinformatics in studying miRNAs in the context of human cancer.

Keywords

microRNAs; miRNAs; therapy; bioinformatics; computational model; pancreatic cancer

1 Introduction

Cancer is one of the deadliest diseases with very low survival rate in the world. It is characterized by an uncontrolled growth of damaged cells. Scientists have been trying to decipher the molecular mechanism of cancer cell formation and the role of onco (cancer promoting) and tumor suppressor (cancer preventing) genes in cancer development [1]. Despite numerous efforts, the cancer cell formation mechanism is yet to be discovered. The discovery of various oncogenes and tumor suppressor genes has provided insight into the biology of cancer and the development of drugs to combat these potential targets [2]. The small non-coding RNAs including miRNAs have shown the potential to act as biomarkers for cancer diagnosis as well as therapeutic agents to cure cancer [3].

*Corresponding Author (+01-402-554-4899; dkbastola@unomaha.edu).

miRNAs are tiny non-coding RNAs that post-transcriptionally regulate the expression of target genes by translational repression or mRNA cleavage. Recently, researchers have observed the role of miRNAs in apoptosis and cell proliferation that are key biological processes in cancer progression and metastasis [4]. The potential role of miRNAs in human cancer diagnosis, progression and metastasis has been studied using various molecular techniques like northern blots, microarray analysis and RNAseq. These miRNAs are expected to provide new insight in cancer research. Recently, the potential role of miRNAs as therapeutic agents has been explored in various cancer types. The oncogenic or tumor suppressor behavior of miRNAs is being exploited to treat cancer. miRNAs can be used as therapeutic agents by either reducing or increasing their expression level or interfering with the miRNA-mRNA regulatory interaction.

Bioinformatics provides a new avenue of understanding cancer biology through intelligent systems. The NIH working definition of Bioinformatics is “Research, development, or application of computational tools and approaches for expanding the use of biological, medical, behavioral or health data, including those to acquire, store, organize, archive, analyze, or visualize such data” [91].

Bioinformatics approaches have shown considerable potential in biomedical research. Computational approaches reduce the search space and provide probabilistic and biologically meaningful outcomes. We consider the systems biology view to be the best path for diagnosing and developing therapies against cancer [5]. Integrating the existing cancer biology knowledge with powerful computational and statistical methods has shown the potential to identify miRNAs as novel biomarkers to diagnose cancer and its various sub-types [6]. Integrating gene and miRNA expression data with computational analysis tools has helped to identify the role of miRNAs in cancer progression and metastasis and their potential role in acting as therapeutic agents in the treatment and cure for cancer [7].

Our goal here is to summarize various existing computational approaches and potential use of bioinformatics in the field of cancer biology. In Section 2, we describe miRNA biogenesis and the mechanism of miRNA mediated post-transcriptional regulation. In section 3, we summarize the role of miRNAs in human cancer. Here, we briefly described the experimental efforts that have been taken to establish the role of miRNA in cancer. In Section 4, we summarize the role of miRNAs as oncogenes and tumor suppressors in various human cancers by providing specific examples from experimental studies. Section 5, we focus on the role of bioinformatics to identify novel miRNAs, their targets and involvement in cellular pathways. In Section 6, we summarize the computational studies done to establish the role of miRNA as therapeutic agent with pancreatic cancer as a case study. Finally, we conclude with the potential of miRNAs as therapeutic agents in human cancers.

2 miRNA Biogenesis

Eukaryotic gene regulation is a complex process involving multiple factors such as transcription machinery, activators, repressors and chromatin. Chromatin maintains inactive genes by guarding them against access by RNA polymerase and other factors. The study of

eukaryotic gene regulation has repeatedly shown the regulatory role of the 5'-end of the gene during transcription. Both enhancer and repressor transcription factors can enhance or decrease gene expression through their interaction at the 5'-end of the gene. Beyond chromatin and transcription factors, the discovery of RNA interference (RNAi) has added a layer to our understanding of gene regulation and the role of non-coding RNA sequences in gene regulation [8, 9].

miRNAs are short ~21–22 nucleotide long non-coding RNAs that have been widely studied as regulators of gene expression [10–12]. In 1993, the first miRNA, lin-4, was identified in *C. elegans* [13, 14]. miRNAs have been found in both protein coding, intronic and intergenic regions. While the miRNAs located in intronic and protein-coding regions are expressed along with their host mRNAs, those found in intergenic regions use their own promoter elements for expression [15]. Interestingly, prior to the discovery of miRNAs, Mizuno *et al*. (1984) showed that translation could be repressed by small RNA (~100 nucleotides) in *E. coli* [16]. Later, these and other studies helped catalyze the discovery of the RNAi process for which Andrew Z. Fire and Craig C. Mello received the Nobel Prize in Physiology in 2006 [17].

The miRNA genes are known to be transcribed in the nucleus by RNA polymerase II or RNA polymerase III into primary miRNA transcripts called pri-miRNAs [18, 19]. As shown in Figure 1, the pri-miRNA is subsequently processed into mature miRNA through cleavage of pri-miRNA by the endonuclease RNA III enzymes — Drosha and Dicer. Cleavage of pri-miRNA in the nucleus by Drosha produces an approximately seventy nucleotide long pre-miRNA [20]. This pre-miRNA is then exported to the cytoplasm where Dicer cleaves pre-miRNA into a 22 nucleotide long duplex containing the mature miRNA (the guide strand) and its antisense compliment (the passenger strand). Gene silencing is achieved through the RNA-induced silencing complex (RISC), an effector ribonucleo-protein complex. RISC is a powerful gene silencing machine controlling gene expression. Pratt and MacRae (2009) have previously reviewed the composition and role of RISC in controlling gene expression [21]. In general, only the guide strand (which has loose pairing at the 5' end) survives within RISC, while the passenger strand is preferentially degraded [8, 22]. A guide strand of the miRNA duplex is incorporated into RISC [23, 24]. RISC identifies target mRNA based on complementarity between the guide miRNA and the mRNA and results in either cleavage of targeted mRNA or translational repression [25, 26].

The miRNAs are endogenous and evolutionarily conserved across the eukaryotic genomes. They are usually clustered on the chromosome [17]. The co-expressed or co-located miRNAs have significance in controlling either same set of target genes or set of target genes with similar biological function. More than 50% of miRNA genes are located in or near cancer-associated genomic regions that represent same chromosomal locations [18]. For example, miR-15a and miR-16a genes, involved in B cell lymphocytic leukemia are both located on chromosome 13 (13q14) [19].

miRNAs act as post-transcriptional gene regulators by generally binding to the 3'-untranslated region (UTR) of their target mRNA. The Watson-Crick base pairing between miRNA and its target sequence results either in the cleavage of the double stranded mRNA

sequence, or translational repression (see Figure 1). There have been some examples of miRNA binding to the either 5'-UTR or coding region of mRNA as well. But the significant binding has been reported in the 3'-UTR region of the mRNA [14, 28].

The untranslated regions (UTRs) of mRNA play a significant role in controlling the behavior of the gene. UTR controls the translation process and its efficiency, stabilizes the mRNA molecule and is involved in subcellular localization as well [29].. Additionally, the 5'-UTRs are known to have sequence- and structure- based motifs that control the translation process and its efficiency. The sequence-based motifs known to be present at the 5'-end of the mRNA include iron-responsive elements (IREs). Others include the internal ribosome entry sites (IRESs) and motifs like the upstream open reading frame (uORFs) and initiation codons. A significant amount (15–50%) of 5' UTRs are known to have uORFs. Furthermore, the 5' cap or 3' poly-A tail present in 5' and 3' UTR respectively, are known to be involved in the stabilization of the mRNA molecule. The 3'-UTR is also capable of mRNA localization [29]. Oftentimes, the mRNAs are localized while attached to the translational machinery. This ensures that translation will occur while improving the efficiency of the process.

Although a single miRNA can bind to the 3'-UTR of multiple mRNA sequences, and a single mRNA can be regulated by multiple miRNAs (i.e., many-to-many relations between miRNA and mRNA), not all miRNA-mRNA interactions have biological relevance [9]. Therefore, identifying relevant regulatory interaction between miRNA and mRNA is critical to our understanding of the regulatory role of miRNAs in cellular pathways and to determining their roles in disease processes.

3 miRNAs and Human cancer

Cancer is a complex polygenic disease caused by the amplification of oncogenes and the mutation of tumor suppressor genes, leading to deregulation of cell proliferation and apoptosis. Hence, understanding gene regulation in the context of cancer is crucial to develop new therapies. The Cancer Genome Project and the Cancer Genome Atlas research efforts represent ongoing attempts to understand the mechanism that underpin tumor formation and progression. Importantly, miRNAs are now known to play a role in cancer formation, invasion and metastasis. Expression studies demonstrate that miRNA genes are deregulated in cancer cells and tissues. Furthermore, more than 50% of miRNA genes have been shown to be located in cancer-associated genomic regions or in fragile sites, suggesting that miRNAs may play an important role in the pathogenesis of human cancers [30].

3.1 miRNAs and cancer —literature mining

The experimental efforts have identified the role of some of the miRNAs in various human cancer types. We performed a computational text mining on the articles related to miRNAs in cancer which have been published in PubMed. Pubmed2Ensembl tool was used to perform text mining of the PubMed database, which is a publicly available resource that consists of more than 23 million citations for biomedical literature from MEDLINE, life science journals, and online books [31]. This tool searched for the keywords in the abstracts of scientific articles in PubMed, then output a list of genes associated with those keywords.

Finally it used Biomart to extract the gene names and other gene-related information [31]. Table 1 summarizes the findings of the text mining. This table also presents information regarding various cancer types and the miRNAs known to be involved in each of these cancer types, including a number of known miRNA target genes. It is evident from the result of the citation count and number of miRNA target genes that the prostate, lung and pancreatic cancer are the most studied cancer types after breast cancer. Numerous efforts are being made to identify the role of miRNAs in cancer and their target genes, which can act as potential therapeutic agents.

3.2 miRNAs as oncogenes

The miRNAs that are up-regulated in cancer can potentially act as oncogenes and are thus referred to as “oncomirs”. Oncomirs negatively regulate tumor suppressor genes leading to uncontrolled cell proliferation. Medina and Slack reviewed potential oncogenic miRNAs and showed that they can play a similar role in several cancer types [32]. A single miRNA can act as an oncogene, such as miR-155, which has been implicated in various hematopoietic malignancies, lung, pancreatic, and breast cancers [33]. Another example of miRNA is the miR-17–92 cluster, whose expression was found to be significantly increased in various lymphomas and lung cancers [34]. Notably, the tumor suppressor gene PTEN (involved in apoptosis) and RB2 have been computationally predicted as targets of miR-17–92 [35]. Additionally, two recent studies have identified miR-19 as the key oncogenic component of miR-17–92, demonstrating that miR-19 inhibits c-Myc-induced apoptosis and promotes c-Myc-mediated lymphomagenesis by repressing the expression of PTEN tumor suppressor gene [34, 36]. The miR-372 and miR-373 miRNAs are examples of oncogenic miRNAs thought to be involved in human testicular germ cell tumors through the inhibition of LATS2 gene expression [37]. Clearly, there are many examples of miRNAs acting as oncogenes, demonstrating their importance to cancer diagnostics and progression.

3.3 miRNAs as tumor suppressors

The decreased expression of miRNA genes in cancer cells leads to another category of miRNAs that can act as tumor suppressors. These miRNAs prevent tumor development by inhibiting oncogenes or genes involved in cell proliferation and apoptosis. The best example of a tumor suppressor miRNA is let-7. Work by Takamizawa *et al*. demonstrated poor expression of let-7 in a large cohort of lung cancer patients [38]. Furthermore, their work confirmed the role of let-7 as a tumor suppressor by over-expressing let-7 miRNA in A549 lung adenocarcinoma cells, which inhibited lung cancer cell growth *in vitro*. The let-7 miRNA has also been reported to negatively regulate the RAS oncogene by binding to the 3′ UTR of RAS mRNA and inducing translational repression [39]. The expression profiling of lung tumor tissues showed significant reduction in levels of let-7 and also showed increased RAS protein levels relative to normal lung tissue, suggesting the role of let-7 as a tumor suppressor gene in lung oncogenesis [39].

4 Role of miRNAs as therapeutics

Given the evidence that miRNAs play a vital role as oncogenes or tumor suppressors, there are numerous efforts to develop new therapies based on miRNA activity. Several strategies

have been designed to manipulate miRNA activity by targeting miRNAs at two different levels: reducing/increasing their expression levels and/or interfering with the miRNA/mRNA interaction.

Restoring the expression of tumor suppressor miRNAs is a potential therapeutic approach. Johnson *et al*. demonstrated that over-expression of the let-7 tumor suppressor miRNA inhibited the growth of lung cancer cells [39]. Also, increasing the expression of miR-26a inhibited tumor progression in an animal model of hepato-cellular carcinoma [40]. Conversely, miRNAs could be synthesized to down-regulate oncogenes and prevent the formation of cancer. Multiple approaches have been designed to down-regulate oncogenic miRNAs. These approaches include the use of anti-miRNA oligonucleotides (AMOs), small molecule inhibitors, miRNA sponges and miRNA masking. AMOs bind to the target mRNA and blocks the miRNA interaction site. *In vivo* studies have been performed to demonstrate the inhibition of MCF-7 cells by designing AMOs against miR-21. Alternatively, artificial miRNAs can be designed to inhibit the expression of oncogenes. In a mouse model, expressing miR-17–92 in transgenic mice strongly inhibited c-myc-induced apoptosis, and resulted in accelerated tumor development [41]. Finally, antisenseRNAs could be used to restrict oncomirs. Recently, researchers had transfected 2′-O-methyl-modified antisense RNA into different miRNAs and showed sequence-specific inhibition [42, 43]. Similarly, modified antisense RNAs known as “antagomirs” have been used to inhibit miRNAs in adult mice [44].

5 Computational study of post-transcriptional regulation by miRNAs

After the discovery of miRNA, researchers tried to identify miRNAs based on the sequence, structure and thermodynamic information in the nucleic acid sequence data. Both experimental and computational approaches have been applied to identify miRNAs. As the experimental identification is a time consuming and resource intensive process, this has led researchers to use computational prediction of miRNAs based on genomic sequence information. A detailed review of computational tools in miRNA discovery was previously done by Gomes *et al*. [45]. The section below provides a brief summary of computational tools categorized into four major computational approaches that have been used in the identification of miRNAs [46–49].

5.1 Computational approaches to predict miRNAs

5.1.1 Sequence- or structure-based approaches—As noted above, the earliest computational approaches to identify miRNAs were based only on sequence or structure conservation. This proved to be a powerful approach to predict miRNAs [50, 51]. The miRNA gene identification primarily focused on locating the origin of miRNA in a genome. The next-generation sequencing (NGS) technology now allows less abundant miRNAs to be identified as well [52]. This requires a use of strong and robust prediction algorithm to identify miRNAs from NGS data. The use of cross-species conservation of sequence and structure has been more recently adopted in many prediction algorithms to identify miRNAs. MiRScan, miRseeker and srnaloop use secondary structure of RNA and conserved stem-loop structure across closely related species in miRNA prediction. The MiRscan identifies RNA secondary structure and then looks for conservation across species, whereas

miRseeker identifies hairpin structures in the conserved regions of RNA sequences. However, many non-conserved miRNAs were missed, even among related species. The lack of sufficient data from related species is another weakness of this comparative approach.

5.1.2 Machine learning-based approaches—A number of machine learning tools such as Support Vector Machines, neural networks, Hidden Markov Model and Naive Bayesian techniques have been widely used to predict miRNAs.

The machine learning methods group similar elements based on their features or attributes. These are artificial intelligence based methods which learn from training data and make predictions based on past observations. For example, to distinguish between normal and cancer patients machine learning system has to learn by training itself on cancer data attributes. Machine learning approach has been successfully used in various other applications outside of the life sciences domain such as spam filtering, face recognition and fraud detection.

The machine learning based computational tools which identify miRNAs using features such as sequence conservation, structure, and folding energy of sequences as their training data. An early example of this approach attempted to identify pre-miRNAs and to predict unknown miRNAs using positive and negative datasets during model training [46–49]. The MiRFinder is an example of computational tool that uses Support Vector Machine with 18 parameters to predict pre-miRNAs [53]. A major limitation of this approach is in the construction of a reliable negative dataset, which can potentially affect the overall efficiency of the prediction method.

5.1.3 Expression data-based approaches—With the advancement of the microarray and the RNAseq technologies, many gene expression studies have been performed in recent years. Gene expression data has shown significant promise in unveiling gene behavior under varying biological conditions. The mRNA or gene expression data has been intensively studied to identify the effect of miRNAs under multiple biological conditions [54].

Research has previously focused primarily on mRNA expression data to develop more robust miRNA identification methods. MiRDeep used a probabilistic model of miRNA biogenesis, which calculated a score of compatibility of position and frequency of sequenced RNA with the secondary structure of precursor miRNA [55]. The score predicted whether the detected RNA was a mature miRNA. By analyzing deep-sequencing data of small RNA molecules, another tool (MiRanalyzer) identified novel miRNAs using a random forest-based machine learning technique. Compared to sequence-based approaches, the use of gene expression data helps to increase the chances to identify more novel miRNAs. However, the major limitation of using gene expression data alone with stringent fold-change threshold is in potentially excluding the genes that are expressed at low levels. The fold change is the measure of amount of gene differentially expressed in a given biological condition.

5.1.4 Integrated approaches—Integrating miRNA and target mRNA expression information has proved to be a valuable approach to miRNA identification. Chang *et al*. developed a reverse prediction tool to predict novel miRNAs [56]. The tool identified

mRNAs expressed at low levels in microarray and expressed sequence tag data. The result was then used to identify the 7-mer motif in the 3'-UTR of the mRNAs. The 7-mer motif is the seed sequence in a mature miRNA that complimentarily binds to the 3'-UTR of mRNA. The 7-mer identified using the tool developed by Chang *et al*. was then compared to existing human miRNA sequences to identify novel miRNAs. As a further advancement to this approach, van Dongen and his colleagues used a ranked gene list to identify the mRNA 7-mer motif [57]. The ranked gene list was obtained by the differential gene expression analysis and the miRNA interference experiments, where the genes identified as highly up or down regulated were ranked based on their p-value [57]. Table 2 summarizes the existing computational tools to predict miRNA genes and their main characteristics.

5.2 Computational approaches to predict miRNA target genes

To decipher the mechanism of miRNA-triggered gene regulation, several attempts have been made to identify potential mRNA targets based on prior knowledge of interaction between the miRNA and its target gene [58]. Web-based solutions that are developed allow researchers to search for targets for a given set of miRNAs. However, web-based approaches inherently limit the tool's use to process high throughput data. The tool such as the miRanda aligns miRNA and mRNA sequences, filters the high scoring pairs, and checks for the thermodynamic stability of potential target site [59]. Alternatively, TargetScan, another web-based tool, uses cross species conservation of target site sequence to predict the mRNAs that are potential miRNA targets [60].

These computational approaches mainly focus on the presence of complementary sequences in the 5'-end of the miRNA and the 3'-UTR of the target mRNA. In plants, target site prediction is based on a perfect complement of the 5'-end of miRNA and 3'-UTR of the target gene. Some of the computational tools for target site prediction in plants include findmiRNA, miRCheck and PatScan [61–63]. Animal miRNAs, on the other hand, bind to their targets by as little as 6–8 nucleotides. These binding sites are also referred as the seed regions. Thus, an alternative to the presence of perfect complementary sequence-based approach is the seed sequence matching. The tool called PicTar uses this seed-sequence based approach and filters the potential target sites on the basis of free energy calculation and species conservation [64]. The software is not available for download. However, it has value for use in miRNA target prediction in organisms other than plants.

A neural network-based target prediction tool, MTar, first used Smith-Waterman alignment to align miRNA to mRNA sequences followed by artificial neural network (ANN) classification into three different target candidate sets [65]. The major limitation of these sequence-based methods is in the loss of potential targets. To overcome this limitation, expression profiles have been integrated together with the sequence complement approach. Integrating expression profile information may be useful in identifying miRNA targets missed by the sequence complement approach [66]. A method that integrates both expression profiles and sequence information was developed by Joung and Fei [67]. In this method, expression values are a feature vector for a Support Vector Machine classifier. Target sites are classified based on expression values and the alignment score between a miRNA-mRNA pair [67]. Overall, both mRNA and miRNA expression profiles can improve

the target identification [68]. Table 3 summarizes the existing computational tools to predict miRNA target genes and their main characteristics.

6 *Insilico* approaches to study the role of miRNA in human cancer

A multitude of factors contribute to the poor diagnosis and treatment of cancer, including severity of disease, resistance to drugs, tumor location and variations in cancer types. Probing cancer mechanisms from a systems biology perspective promises a better understanding of the disease. Hence, it is important to integrate systems biology, cancer research and bioinformatics to gain a more complete and accurate picture of cancer. This integrated approach will lead to advancements in the diagnosis, prognosis and treatment of cancer.

Cancer bioinformatics is an emergent field that integrates the existing knowledge from cancer biologists, information technology and mathematics. Its goal is to develop bioinformatics models and tools to answer cancer-related questions. Existing cancer therapies are often only effective in a subset of cancer patients. Hence, personalized cancer therapy is a promising alternate approach. Integrating different omics research efforts is expected to lead to a customized therapy based on the subtype, severity and sensitivity of the cancer in a patient.

Recent reports have shown the potential of integrating bioinformatics tools in answering cancer-specific questions. *In silico* work can act as an exploratory tool to reduce the search space and guide cancer biologists to perform focused experiments. This view is supported by Materi *et al*., who indicated that computational systems modeling can be a useful tool for biologists if they produce biologically sound results which facilitate experiments [69]. Computational modeling can assist biologists to run experiments *in silico* to conserve resources and speed the process of cancer diagnosis and therapies.

With high-throughput technologies, tremendous volumes of experimental data are being generated. Managing and maintaining these enormous amounts of data is a persistent issue. Computer science has contributed greatly to the field of data management. Multiple web-based databases have been developed for the storage and retrieval of genome-scale data including genomic sequences, genome annotations, promoter sequences, transcriptomics, proteomics and structural information [7]. Databases help in efficiently maintaining, distributing and accessing experimental data. This has enabled data sharing among peers and the integration of knowledge to address some of the fundamental issues in cancer research.

6.1 Computational models to identify miRNA-target pairs

Computational approaches help identify miRNAs and their potential targets using features such as sequence complementation, miRNA-mRNA duplex structure, binding energy between miRNA-mRNA duplex and properties of flanking regions around miRNA-mRNA interaction sites. In our previous work, we developed a framework to identify potential targets of miRNAs by optimizing and using the miRNA-mRNA relationship [70]. Our approach is generic and can be applied to identify a potential set of miRNA-mRNA pairs involved in various human cancer types. This smaller potential set of target genes can be

later validated experimentally, thereby reducing the resources required to validate all potential miRNA targets in a wet-lab setting.

Several other attempts have been made to identify miRNA-mRNA pairs based on prior knowledge of interaction that exist between the miRNA and its target genes. Some of the previously described miRNA target prediction tools in Section 5.2 can be used to identify miRNA-mRNA pairs. There have been attempts to integrate a number of different genomic and proteomic data to better predict the miRNA-mRNA relationships [52, 53, 71–76]. The methods based on sequence complementarity and integrating miRNA and mRNA expression data to identify miRNA-mRNA relationship are mainly statistical models, which provide a significance value for a miRNA-mRNA pair to be a potential expression regulation site. However, due to heterogeneous nature of genomic and proteomic data, such integration is a big challenge and is expected to lead to high false positive results.

6.2 Differential gene expression analysis

Microarray expression profiling is a common approach to identify up- and down-regulated genes and miRNAs in cancer cells. The identification of deregulated miRNAs helps to identify the role of miRNA in the cancer cell. For instance, differential analysis of miRNA gene expression levels in normal versus cancer cells/tissues can yield clues as to whether the miRNA is likely to act as an oncomir or tumor suppressor. Bioinformatics plays an important role in designing robust and efficient statistical models to analyze the miRNA expression data. Expression data obtained using microarray techniques suffer from missing and low quality data problems due to experimental artifacts. Analyzing this low quality data might lead to inaccurate conclusions. More robust analysis models are required to mitigate the effect of errors that can occur during microarray experiments and to ensure the reliability of the results. Mathematical approaches, such as background correction and normalization, can be applied to raw expression data to increase the robustness of data as well as to accurately analyze the expression profiling data. The deterministic or probabilistic models model the progression of tumor in a cell or tissue and classify different cancer types and sub types. The deterministic model requires the knowledge of data beforehand whereas probabilistic models are based on probability of the event to occur based on the observations. With limited knowledge about the different cancers and their sub-types, it is very challenging to predict each cancer type and its subtype with high accuracy. The probabilistic methods have been very productive in differentiating cancer subtypes with comfortable accuracy values.

Both supervised and un-supervised classifiers have been developed to classify the cancer types and sub-types accurately. The supervised classifier models based on support vector machine, artificial neural networks, and random forest have been developed to address the cancer sub-type classification problem [70–75]. The unsupervised classifiers based on hierarchical clustering and k-means clustering have also been developed to produce accurate classification results with limited knowledge about the cancer sub-types [7]. Table 4 summarizes the existing computational approaches that use the gene expression data in cancer diagnostics, classifying cancer subtypes and in functional enrichment of differentially expressed genes in various tumor conditions.

6.3 Pathway enrichment models

Pathway enrichment analysis is the process of mapping genes identified by high-throughput technologies to known pathways or Gene Ontology terms [76]. The human genome project has made substantial contributions toward annotating human genes and the associations among genes. Curated databases like KEGG, Reactome and Gene Ontology have been developed to categorize genes into functional modules [76–79]. Pathway enrichment of genes identifies enriched biological processes and cancer pathways. Choosing specific and selective miRNA target(s) is a crucial step in using miRNA as a therapeutic agent in cancer treatment. Computational approaches can help to identify the pool of cancer-related biological processes that involve genes targeted by a therapeutic miRNA. This can in turn help in targeting those miRNA target genes that are involved in cancer pathways but not in cellular processes that are essential for the survival of the cell. Table 4 summarizes the existing computational approaches that apply pathway enrichment approach to identify pathways and biological processes involved in cancer. It also summarizes the tools that use gene co-expression data to identify similarly behaving genes in tumor conditions.

6.4 Case Study: Pancreatic cancer and miRNAs

Pancreatic cancer, or pancreatic adenocarcinoma, is the most common cause of cancer-related deaths in the United States. The current 5-year survival rate for pancreatic cancer patients treated with state-of-the art therapies is 5% [93]. The high mortality rate of pancreatic cancer patients is due in large part because of the inability to detect pancreatic cancer in its early stages. Patients diagnosed with advanced stage pancreatic cancer are not candidates for surgical resection, and respond poorly to chemotherapy and radiation. Thus, research has focused on identifying minimally invasive diagnostic markers. Currently, the CA19-9 tumor marker is the most reliable diagnostic serum marker for pancreatic cancer; however, even CA19-9 is limited in its ability to detect early/small tumors [94]. Given that several miRNAs are deregulated in pancreatic cancer tissues, their utility as potential biomarkers is currently being explored. Wang *et al*. summarized a list of miRNAs identified as up- and down-regulated in several pancreatic cancer tissue expression studies. Table 5 is the updated version of table published by [95].

The expression patterns of miRNAs can help to distinguish pancreatic cancer cells from normal tissues. Indeed, researchers were successfully able to distinguish between pancreatic cancer, chronic pancreatitis and normal pancreas [101]. Furthermore, a number of studies suggest that miR-155, miR-203, miR-210 and miR-222 are associated with poor survival of pancreatic cancer patients [96, 102, 103]. Although there seems to be a trend to identify up-regulated miRNAs in pancreatic cancer (as shown in Table 5), but down-regulated miRNAs can also serve as potential biomarkers in pancreatic cancer diagnosis. Expression of the down-regulated miRNAs could be restored through construction of artificial miRNAs, which may in turn inhibit oncogenes. Overall, such expression studies clearly indicate that miRNAs are functionally involved in the underlying mechanisms of pancreatic cancer development and progression. Ongoing attempts to target these deregulated miRNAs through antisense oligonucleotides or artificial miRNAs hold promise for pancreatic cancer.

Bioinformatics continues to play an important role in the study of pancreatic cancer. Computational approaches have helped biologists to rapidly identify potential biomarkers and design therapies to treat pancreatic cancer. The focus must now be turned to expediting the detection and diagnosis of early stage pancreatic cancer. Bioinformatics approaches have been developed to analyze gene expression profiles in pancreatic cancer to identify the role of miRNAs in early diagnosis and treatment [104].Similar strategy was developed to analyze gene expression profile in chronic pancreatitis and identify the role of miRNAs in early diagnosis and treatment [105]. The gene expression data was downloaded from Gene Expression Omnibus and differential gene expression analysis was performed to identify differentially expressed genes in mice. The miRNAs targeting the differentially regulated genes were identified based on regulatory relationship between miRNAs and genes. The authors identified miR-124a as a potential target for the diagnosis and treatment of chronic pancreatitis [105]. The differential gene expression studies were also able to distinguish between pancreatic adenocarcinoma from normal and chronic pancreatitis [106]. The statistical analysis was performed to identify the most differentially regulated miRNAs in pancreatic adenocarcinoma samples.

The computational studies can also identify potential biological pathways involved in pancreatic cancer. Various approaches have been developed to identify pathways, especially immune-related, involved in pancreatic cancer [107, 108]. These research efforts include the use of gene expression profiling data to identify dysregulated genes in pancreatic cancer and then performing functional enrichment of differentially expressed genes. The differentially expressed genes were found to be highly enriched in immune related pathways. This implies that dysregulated pathways in pancreatic cancer are associated with the immune system, which can provide useful information for potential treatment of pancreatic cancer [107]. Apart from identification of immune related pathways involved in pancreatic cancer, researchers have also identified activation of mTOR signaling pathway and renal cell carcinoma pathway to be associated with pancreatic cancer [108].As described earlier, these computational efforts can help researchers to identify potential miRNA targets by analyzing miRNA and mRNA expression data. Such approaches can greatly reduce the search space allowing molecular biologists to validate a much smaller pool of miRNAs which can serve as true biomarkers.

Systems biology approach to study the significant contribution of miRNA in cancer biology can help research community to decipher the role of miRNA in regulating onco-or tumor suppressor genes. Network-based studies have also been performed to identify miRNAs that can influence cancer. Previous work has developed a miRNA network that influence cancer [86]. In these studies, statistical models were used to identify miRNAs that play a significant role in cancer. An integrated framework was developed to infer gene co-expression networks by integrating statistical models and visualization tools [5, 109]. Also, the network-based strategies were applied in the analysis of two immune related pathways: the B-cell receptor and Nod-like receptor signaling pathways [86]. The gene sub-network identified by this strategy offers new insights and potential targets for development of prognostic and therapeutic treatments. Similar approaches can certainly be applied and will likely benefit the study of pancreatic cancer [85, 86, 95, 110].

7 Conclusions

In this review, we have discussed the potential role of miRNA in cancer propagation and its use as a therapeutic agent. Understanding the detailed mechanism behind miRNA acting as oncomiRs or tumor suppressors has not yet been achieved. It is estimated that human genome has more miRNAs to be discovered. The functional association of known miRNAs is still an open-ended question. Despite the relentless efforts by cancer biologists to predict miRNAs and its potential mRNA targets, the miRNAs has been seldom used as biomarkers to diagnose cancers or as therapeutic agents to treat cancer patients. Experimentally validating the role of every human miRNA as a potential biomarker or therapeutic agent is very resource-intensive and time-consuming. Innovative multi-disciplinary approaches need to be designed to achieve rapid progress in cancer therapeutics. Bioinformatics can be considered as an integrated, multidisciplinary science that has the potential to deliver new ways of effectively treat cancer patients. The use of computational approaches not only speeds up the process in identifying miRNAs with a potential role in cancer biology but also reduces the search space to more likely miRNAs that can be experimentally validated by biologists in a lab.

Acknowledgements

We would like to thank INBRE grant number P20 RR16469 from the National Center for Research Resources (NCRR) a component of the National Institute of Health (NIH) for their support. The content of this publication is the sole responsibility of the authors and does not necessarily represent the official views of NCRR or NIH.

References

1. Abba M, Mudduluru G, Allgayer H. MicroRNAs in cancer: small molecules, big chances. Anticancer Agents Med. Chem. 2012; 12:733–743. [PubMed: 22292749]
2. Broderick JA, Zamore PD. MicroRNA therapeutics. Gene Ther. 2011; 18:1104–1110. [PubMed: 21525952]
3. Iorio MV, Croce CM. MicroRNA dysregulation in cancer: diagnostics, monitoring and therapeutics. A comprehensive review. EMBO Mol. Med. 2012; 4:143–159. [PubMed: 22351564]
4. Ambros V. The functions of animal microRNAs. Nature. 2004; 431:350–355. [PubMed: 15372042]
5. Blair RH, Trichler DL, Gaille DP. Mathematical and statistical modeling in cancer systems biology. Front. Physiol. 2012; 3:227. [PubMed: 22754537]
6. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, Downing JR, Jacks T, Horvitz HR, Golub TR. MicroRNA expression profiles classify human cancers. Nature. 2005; 435:834–838. [PubMed: 15944708]
7. Edelman LB, Eddy JA, Price ND. In silico models of cancer. Wiley Interdiscip. Rev. Syst. Biol. Med. 2010; 2:438–459. [PubMed: 20836040]
8. Schwarz DS, Hutvagner G, Du T, Xu Z, Aronin N, Zamore PD. Asymmetry in the assembly of the RNAi enzyme complex. Cell. 2003; 115:199–208. [PubMed: 14567917]
9. Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell. 2009; 136:215–233. [PubMed: 19167326]
10. Alvarez-Garcia I, Miska EA. MicroRNA functions in animal development and human disease. Development. 2005; 132:4653–4662. [PubMed: 16224045]
11. Cho WC. OncomiRs: the discovery and progress of microRNAs in cancers. Mol Cancer. 2007; 6:60. [PubMed: 17894887]
12. Felicetti F, Errico MC, Bottero L, Segnalini P, Stoppacciaro A, Biffoni M, Felli N, Mattia G, Petrini M, Colombo MP, Peschle C, Care A. The promyelocytic leukemia zinc finger-

microRNA-221/-222 pathway controls melanoma progression through multiple oncogenic mechanisms. Cancer Res. 2008; 68:2745–2754. [PubMed: 18417445]

13. Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell. 1993; 75:843–854. [PubMed: 8252621]
14. Wightman B, Ha I, Ruvkun G. Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell. 1993; 75:855–862. [PubMed: 8252622]
15. Baskerville S, Bartel DP. Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. RNA. 2005; 11:241–247. [PubMed: 15701730]
16. Mizuno T, Chou MY, Inouye M. A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). Proc. Natl. Acad. Sci. U. S. A. 1984; 81:1966–1970. [PubMed: 6201848]
17. Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC. Potent and specific genetic interference by double-stranded RNA in Caenorhabditis elegans. Nature. 1998; 391:806–811. [PubMed: 9486653]
18. Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN. MicroRNA genes are transcribed by RNA polymerase II. EMBO J. 2004; 23:4051–4060. [PubMed: 15372072]
19. Borchert GM, Lanier W, Davidson BL. RNA polymerase III transcribes human microRNAs. Nat. Struct. Mol. Biol. 2006; 13:1097–1101. [PubMed: 17099701]
20. Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Radmark O, Kim S, Kim VN. The nuclear RNase III Drosha initiates microRNA processing. Nature. 2003; 425:415–419. [PubMed: 14508493]
21. Pratt AJ, MacRae IJ. The RNA-induced silencing complex: a versatile gene-silencing machine. J. Biol. Chem. 2009; 284:17897–17901. [PubMed: 19342379]
22. Khvorova A, Reynolds A, Jayasena SD. Functional siRNAs and miRNAs exhibit strand bias. Cell. 2003; 115:209–216. [PubMed: 14567918]
23. Hammond SM, Bernstein E, Beach D, Hannon GJ. An RNA-directed nuclease mediates post-transcriptional gene silencing in Drosophila cells. Nature. 2000; 404:293–296. [PubMed: 10749213]
24. Martinez J, Patkaniowska A, Urlaub H, Luhrmann R, Tuschl T. Single-stranded antisense siRNAs guide target RNA cleavage in RNAi. Cell. 2002; 110:563–574. [PubMed: 12230974]
25. Bartel DP. MicroRNAs: genomics biogenesis, mechanism, and function. Cell. 2004; 116:281–297. [PubMed: 14744438]
26. Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, Bartel DP, Linsley PS, Johnson JM. Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature. 2005; 433:769–773. [PubMed: 15685193]
27. Banwait JK, Ali HH, Bastola DR. Optimisation of miRNA-mRNA relationship prediction using biological features. Int. J. Comput. Biol. Drug Des. 2014; 7:45–60. [PubMed: 24429502]
28. Lai EC. Micro RNAs are complementary to 3' UTR sequence motifs that mediate negative post-transcriptional regulation. Nat. Genet. 2002; 30:363–364. [PubMed: 11896390]
29. Mignone F, Gissi C, Liuni S, Pesole G. Untranslated regions of mRNAs. Genome Biol. 2002; 3 REVIEWS0004.
30. Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, Shimizu M, Rattan S, Bullrich F, Negrini M, Croce CM. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc. Natl. Acad. Sci. U. S. A. 2004; 101:2999–3004. [PubMed: 14973191]
31. Baran J, Gerner M, Haeussler M, Nenadic G, Bergman CM. Pubmed2ensembl: a Resource for Mining the Biological Literature on Genes. PLoS One. 2011; 6:e24716. [PubMed: 21980353]
32. Medina PP, Nolde M, Slack FJ. OncomiR addiction in an in vivo model of microRNA-21-induced pre-B-cell lymphoma. Nature. 2010; 467:86–90. [PubMed: 20693987]
33. Kluiver J, Kroesen BJ, Poppema S, van den Berg A. The role of microRNAs in normal hematopoiesis and hematopoietic malignancies. Leukemia. 2006; 20:1931–1936. [PubMed: 16990772]

34. Mu P, Han YC, Betel D, Yao E, Squatrito M, Ogrodowski P, de Stanchina E, D'Andrea A, Sander C, Ventura A. Genetic dissection of the miR-17~92 cluster of microRNAs in Myc-induced B-cell lymphomas. Genes Dev. 2009; 23:2806–2811. [PubMed: 20008931]
35. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of mammalian microRNA targets. Cell. 2003; 115:787–798. [PubMed: 14697198]
36. Olive V, Bennett MJ, Walker JC, Ma C, Jiang I, Cordon-Cardo C, Li QJ, Lowe SW, Hannon GJ, He L. miR-19 is a key oncogenic component of mir-17–92. Genes Dev. 2009; 23:2839–2849. [PubMed: 20008935]
37. Voorhoeve PM, le Sage C, Schrier M, Gillis AJ, Stoop H, Nagel R, Liu YP, van Duijse J, Drost J, Griekspoor A, Zlotorynski E, Yabuta N, De Vita G, Nojima H, Looijenga LH, Agami R. A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors. Cell. 2006; 124:1169–1181. [PubMed: 16564011]
38. Takamizawa J, Konishi H, Yanagisawa K, Tomida S, Osada H, Endoh H, Harano T, Yatabe Y, Nagino M, Nimura Y, Mitsudomi T, Takahashi T. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. Cancer Res. 2004; 64:3753–3756. [PubMed: 15172979]
39. Johnson SM, Grosshans H, Shingara J, Byrom M, Jarvis R, Cheng A, Labourier E, Reinert KL, Brown D, Slack FJ. RAS is regulated by the let-7 microRNA family. Cell. 2005; 120:635–647. [PubMed: 15766527]
40. Li W, Lebrun DG, Li M. The expression and functions of microRNAs in pancreatic adenocarcinoma and hepatocellular carcinoma. Chin. J. Cancer. 2011; 30:540–550. [PubMed: 21801602]
41. He L, Thomson JM, Hemann MT, Hernando-Monge E, Mu D, Goodson S, Powers S, Cordon-Cardo C, Lowe SW, Hannon GJ, Hammond SM. A microRNA polycistron as a potential human oncogene. Nature. 2005; 435:828–833. [PubMed: 15944707]
42. Hutvagner G, Simard MJ, Mello CC, Zamore PD. Sequence-specific inhibition of small RNA function. PLoS Biol. 2004; 2:E98. [PubMed: 15024405]
43. Meister G, Landthaler M, Dorsett Y, Tuschl T. Sequence-specific inhibition of microRNA- and siRNA-induced RNA silencing. RNA. 2004; 10:544–550. [PubMed: 14970398]
44. Krutzfeldt J, Rajewsky N, Braich R, Rajeev KG, Tuschl T, Manoharan M, Stoffel M. Silencing of microRNAs in vivo with 'antagomirs'. Nature. 2005; 438:685–689. [PubMed: 16258535]
45. Gomes CP, Cho JH, Hood L, Franco OL, Pereira RW, Wang K. A Review of Computational Tools in microRNA Discovery. Front. Genet. 2013; 4:81. [PubMed: 23720668]
46. Hertel J, Stadler PF. Hairpins in a Haystack: recognizing microRNA precursors in comparative genomics data. Bioinformatics. 2006; 22:e197–e202. [PubMed: 16873472]
47. Pfeffer S, Sewer A, Lagos-Quintana M, Sheridan R, Sander C, Grasser FA, van Dyk LF, Ho CK, Shuman S, Chien M, Russo JJ, Ju J, Randall G, Lindenbach BD, Rice CM, Simon V, Ho DD, Zavolan M, Tuschl T. Identification of microRNAs of the herpesvirus family. Nat. Methods. 2005; 2:269–276. [PubMed: 15782219]
48. Xue C, Li F, He T, Liu GP, Li Y, Zhang X. Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine. BMC Bioinformatics. 2005; 6:310. [PubMed: 16381612]
49. Sewer A, Paul N, Landgraf P, Aravin A, Pfeffer S, Brownstein MJ, Tuschl T, van Nimwegen E, Zavolan M. Identification of clustered microRNAs using an ab initio prediction method. BMC Bioinformatics. 2005; 6:267. [PubMed: 16274478]
50. Lim LP, Lau NC, Weinstein EG, Abdelhakim A, Yekta S, Rhoades MW, Burge CB, Bartel DP. The microRNAs of Caenorhabditis elegans. Genes Dev. 2003; 17:991–1008. [PubMed: 12672692]
51. Wang X, Zhang J, Li F, Gu J, He T, Zhang X, Li Y. MicroRNA identification based on sequence and structure alignment. Bioinformatics. 2005; 21:3610–3614. [PubMed: 15994192]
52. Bar M, Wyman SK, Fritz BR, Qi J, Garg KS, Parkin RK, Kroh EM, Bendoraite A, Mitchell PS, Nelson AM, Ruzzo WL, Ware C, Radich JP, Gentleman R, Ruohola-Baker H, Tewari M. MicroRNA discovery and profiling in human embryonic stem cells by deep sequencing of small RNA libraries. Stem Cells. 2008; 26:2496–2505. [PubMed: 18583537]

53. Huang TH, Fan B, Rothschild MF, Hu ZL, Li K, Zhao SH. MiRFinder: an improved approach and software implementation for genome-wide fast microRNA precursor scans. BMC Bioinformatics. 2007; 8:341. [PubMed: 17868480]

54. Landgraf P, Rusu M, Sheridan R, Sewer A, Iovino N, Aravin A, Pfeffer S, Rice A, Kamphorst AO, Landthaler M, Lin C, Socci ND, Hermida L, Fulci V, Chiaretti S, Foa R, Schliwka J, Fuchs U, Novosel A, Muller RU, Schermer B, Bissels U, Inman J, Phan Q, Chien M, Weir DB, Choksi R, De Vita G, Frezzetti D, Trompeter HI, Hornung V, Teng G, Hartmann G, Palkovits M, Di Lauro R, Wernet P, Macino G, Rogler CE, Nagle JW, Ju J, Papavasiliou FN, Benzing T, Lichter P, Tam W, Brownstein MJ, Bosio A, Borkhardt A, Russo JJ, Sander C, Zavolan M, Tuschl T. A mammalian microRNA expression atlas based on small RNA library sequencing. Cell. 2007; 129:1401–1414. [PubMed: 17604727]

55. Friedlander MR, Chen W, Adamidi C, Maaskola J, Einspanier R, Knespel S, Rajewsky N. Discovering microRNAs from deep sequencing data using miRDeep. Nat. Biotechnol. 2008; 26:407–415. [PubMed: 18392026]

56. Chang YM, Juan HF, Lee TY, Chang YY, Yeh YM, Li WH, Shih AC. Prediction of human miRNAs using tissue-selective motifs in 3' UTRs. Proc. Natl. Acad. Sci. U. S. A. 2008; 105:17061–17066. [PubMed: 18952845]

57. van Dongen S, Abreu-Goodger C, Enright AJ. Detecting microRNA binding and siRNA off-target effects from expression data. Nat. Methods. 2008; 5:1023–1025. [PubMed: 18978784]

58. Lai EC. Predicting and validating microRNA targets. Genome Biol. 2004; 5:115. [PubMed: 15345038]

59. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in Drosophila. Genome Biol. 2003; 5:R1. [PubMed: 14709173]

60. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell. 2005; 120:15–20. [PubMed: 15652477]

61. Adai A, Johnson C, Mlotshwa S, Archer-Evans S, Manocha V, Vance V, Sundaresan V. Computational prediction of miRNAs in Arabidopsis thaliana. Genome Res. 2005; 15:78–91. [PubMed: 15632092]

62. Jones-Rhoades MW, Bartel DP. Computational identification of plant microRNAs their targets, including a stress-induced miRNA. Mol Cell. 2004; 14:787–799. [PubMed: 15200956]

63. Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP. Prediction of plant microRNA targets. Cell. 2002; 110:513–520. [PubMed: 12202040]

64. Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M, Rajewsky N. Combinatorial microRNA target predictions. Nat Genet. 2005; 37:495–500. [PubMed: 15806104]

65. Chandra V, Girijadevi R, Nair AS, Pillai SS, Pillai RM. MTar: a computational microRNA target prediction architecture for human transcriptome. BMC Bioinformatics. 2010; 11(Suppl 1):S2. [PubMed: 20122191]

66. Jones-Rhoades MW, Bartel DP, Bartel B. MicroRNAS and their regulatory roles in plants. Annu Rev Plant Biol. 2006; 57:19–53. [PubMed: 16669754]

67. Joung JG, Fei Z. Computational identification of condition-specific miRNA targets based on gene expression profiles and sequence information. BMC Bioinformatics. 2009; 10(Suppl.1):S34. [PubMed: 19208135]

68. Huang JC, Babak T, Corson TW, Chua G, Khan S, Gallie BL, Hughes TR, Blencowe BJ, Frey BJ, Morris QD. Using expression profiling data to identify human microRNA targets. Nat Methods. 2007; 4:1045–1049. [PubMed: 18026111]

69. Materi W, Wishart DS. Computational systems biology in cancer: modeling methods and applications. Gene Regul. Syst. Bio. 2007; 1:91–110.

70. Chandra B, Gupta M. An efficient statistical feature selection approach for classification of gene expression data. J. Biomed. Inform. 2011; 44:529–535. [PubMed: 21241823]

71. Chen Z, Li J, Wei L. A multiple kernel support vector machine scheme for feature selection and rule extraction from gene expression data of cancer tissue. Artif. Intell. Med. 2007; 41:161–175. [PubMed: 17851055]

72. Chow ML, Moler EJ, Mian IS. Identifying marker genes in transcription profiling data using a mixture of feature relevance experts. Physiol. Genomics. 2001; 5:99–111. [PubMed: 11242594]

73. Maulik U, Mukhopadhyay A, Chakraborty D. Gene-expression-based cancer subtypes prediction through feature selection transductive SVM, IEEE Trans. Biomed. Eng. 2013; 60:1111–1117.

74. Mukhopadhyay A, Bandyopadhyay S, Maulik U. Multi-class clustering of cancer subtypes through SVM based ensemble of pareto-optimal solutions for gene marker identification. PLoS One. 2010; 5:e13803. [PubMed: 21103052]

75. Wang H, Zhang H, Dai Z, Chen MS, Yuan Z. TSG: a new algorithm for binary and multi-class cancer classification and informative genes selection. BMC Med. Genomics. 2013; 6(Suppl. 1) S3-8794-6-S1-S3. Epub 2013 Jan 23.

76. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, Harris MA, Hill DP, Issel-Tarver L, Kasarskis A, Lewis S, Matese JC, Richardson JE, Ringwald M, Rubin GM, Sherlock G. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. 2000; 25:25–29. [PubMed: 10802651]

77. Mi H, Thomas P. PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. Methods Mol. Biol. 2009; 563:123–140. [PubMed: 19597783]

78. Croft D, Mundo AF, Haw R, Milacic M, Weiser J, Wu G, Caudy M, Garapati P, Gillespie M, Kamdar MR, Jassal B, Jupe S, Matthews L, May B, Palatnik S, Rothfels K, Shamovsky V, Song H, Williams M, Birney E, Hermjakob H, Stein L, D'Eustachio P. The Reactome pathway knowledgebase. Nucleic Acids Res. 2014; 42:D472–D477. [PubMed: 24243840]

79. Kanehisa M, Goto S, Kawashima S, Nakaya A. The KEGG databases at GenomeNet. Nucleic Acids Res. 2002; 30:42–46. [PubMed: 11752249]

80. Liu R, Chen X, Du Y, Yao W, Shen L, Wang C, Hu Z, Zhuang R, Ning G, Zhang C, Yuan Y, Li Z, Zen K, Ba Y, Zhang CY. Serum microRNA expression profile as a biomarker in the diagnosis and prognosis of pancreatic cancer. Clin. Chem. 2012; 58:610–618. [PubMed: 22194634]

81. Yeoh EJ, Ross ME, Shurtleff SA, Williams WK, Patel D, Mahfouz R, Behm FG, Raimondi SC, Relling MV, Patel A, Cheng C, Campana D, Wilkins D, Zhou X, Li J, Liu H, Pui CH, Evans WE, Naeve C, Wong L, Downing JR. Classification subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. Cancer. Cell. 2002; 1:133–143. [PubMed: 12086872]

82. Sorlie T, Perou CM, Tibshirani R, Aas T, Geisler S, Johnsen H, Hastie T, Eisen MB, van de Rijn M, Jeffrey SS, Thorsen T, Quist H, Matese JC, Brown PO, Botstein D, Lonning PE, Borresen-Dale AL. Gene expression patterns of breast carcinomas distinguish tumor subclasses with clinical implications. Proc. Natl. Acad. Sci. U. S. A. 2001; 98:10869–10874. [PubMed: 11553815]

83. Alizadeh AA, Eisen MB, Davis RE, Ma C, Lossos IS, Rosenwald A, Boldrick JC, Sabet H, Tran T, Yu X, Powell JI, Yang L, Marti GE, Moore T, Hudson J Jr, Lu L, Lewis DB, Tibshirani R, Sherlock G, Chan WC, Greiner TC, Weisenburger DD, Armitage JO, Warnke R, Levy R, Wilson W, Grever MR, Byrd JC, Botstein D, Brown PO, Staudt LM. Distinct types of diffuse large B-cell lymphoma identified by gene expression profiling. Nature. 2000; 403:503–511. [PubMed: 10676951]

84. Bittner M, Meltzer P, Chen Y, Jiang Y, Seftor E, Hendrix M, Radmacher M, Simon R, Yakhini Z, Ben-Dor A, Sampas N, Dougherty E, Wang E, Marincola F, Gooden C, Lueders J, Glatfelter A, Pollock P, Carpten J, Gillanders E, Leja D, Dietrich K, Beaudry C, Berens M, Alberts D, Sondak V. Molecular classification of cutaneous malignant melanoma by gene expression profiling. Nature. 2000; 406:536–540. [PubMed: 10952317]

85. Hunten S, Siemens H, Kaller M, Hermeking H. The p53/microRNA network in cancer: experimental and bioinformatics approaches. Adv. Exp. Med. Biol. 2013; 774:77–101. [PubMed: 23377969]

86. Zhu M, Xu Z, Wang K, Wang N, Li Y. microRNA and gene networks in human pancreatic cancer. Oncol. Lett. 2013; 6:1133–1139. [PubMed: 24137477]

87. Mansmann U, Jurinovic V. Biological feature validation of estimated gene interaction networks from microarray data: a case study on MYC in lymphomas. Briefings in Bioinformatics. 2011; 12:230–244. [PubMed: 21498549]

88. Gong J, Diao B, Yao GJ, Liu Y, Xu GZ. Analysis of regulatory networks constructed based on gene coexpression in pituitary adenoma. J. Genet. 2013; 92:489–497. [PubMed: 24371170]

89. Clarke C, Madden SF, Doolan P, Aherne ST, Joyce H, O'Driscoll L, Gallagher WM, Hennessy BT, Moriarty M, Crown J, Kennedy S, Clynes M. Correlating transcriptional networks to breast cancer survival: a large-scale coexpression analysis. Carcinogenesis. 2013; 34:2300–2308. [PubMed: 23740839]

90. Penrod NM, Moore JH. Influence networks based on coexpression improve drug target discovery for the development of novel cancer therapeutics. BMC Syst. Biol. 2014; 8 12-0509-8-12.

91. Penrod NM, Moore JH. Key genes for modulating information flow play a temporal role as breast tumor coexpression networks are dynamically rewired by letrozole. BMC Med. Genomics. 2013; 6(Suppl. 2) S2-8794-6-S2-S2. Epub 2013 May 7.

92. Materi W, Wishart DS. Computational systems biology in cancer: modeling methods and applications. Gene Regul. Syst. Bio. 2007; 1:91–110.

93. Siegel R, Naishadham D, Jemal A. Cancer statistics, 2012. CA Cancer. J. Clin. 2012; 62:10–29. [PubMed: 22237781]

94. Harsha HC, Kandasamy K, Ranganathan P, Rani S, Ramabadran S, Gollapudi S, Balakrishnan L, Dwivedi SB, Telikicherla D, Selvan LD, Goel R, Mathivanan S, Marimuthu A, Kashyap M, Vizza RF, Mayer RJ, Decaprio JA, Srivastava S, Hanash SM, Hruban RH, Pandey A. A compendium of potential biomarkers of pancreatic cancer. PLoS Med. 2009; 6:e1000046. [PubMed: 19360088]

95. Wang J, Sen S. MicroRNA functional network in pancreatic cancer: from biology to biomarkers of disease. J. Biosci. 2011; 36:481–491. [PubMed: 21799259]

96. Park JK, Henry JC, Jiang J, Esau C, Gusev Y, Lerner MR, Postier RG, Brackett DJ, Schmittgen TD. miR-132 and miR-212 are increased in pancreatic cancer and target the retinoblastoma tumor suppressor. Biochem. Biophys. Res. Commun. 2011; 406:518–523. [PubMed: 21329664]

97. Morimura R, Komatsu S, Ichikawa D, Takeshita H, Tsujiura M, Nagata H, Konishi H, Shiozaki A, Ikoma H, Okamoto K, Ochiai T, Taniguchi H, Otsuji E. Novel diagnostic value of circulating miR-18a in plasma of patients with pancreatic cancer. Br. J. Cancer. 2011; 105:1733–1740. [PubMed: 22045190]

98. Giovannetti E, Erozenci A, Smit J, Danesi R, Peters GJ. Molecular mechanisms underlying the role of microRNAs (miRNAs) in anticancer drug resistance and implications for clinical practice. Crit. Rev. Oncol. Hematol. 2012; 81:103–122. [PubMed: 21546262]

99. Liu C, Cheng H, Shi S, Cui X, Yang J, Chen L, Cen P, Cai X, Lu Y, Wu C, Yao W, Qin Y, Liu L, Long J, Xu J, Li M, Yu X. MicroRNA-34b inhibits pancreatic cancer metastasis through repressing Smad3. Curr. Mol. Med. 2013; 13:467–478. [PubMed: 23305226]

100. Zhao G, Wang B, Liu Y, Zhang JG, Deng SC, Qin Q, Tian K, Li X, Zhu S, Niu Y, Gong Q, Wang CY. miRNA-141 downregulated in pancreatic cancer, inhibits cell proliferation and invasion by directly targeting MAP4K4. Mol. Cancer. Ther. 2013; 12:2569–2580. [PubMed: 24013097]

101. Bloomston M, Frankel WL, Petrocca F, Volinia S, Alder H, Hagan JP, Liu CG, Bhatt D, Taccioli C, Croce CM. MicroRNA expression patterns to differentiate pancreatic adenocarcinoma from normal pancreas and chronic pancreatitis. JAMA. 2007; 297:1901–1908. [PubMed: 17473300]

102. Hwang JH. Pancreatic cancer and microRNAs. Korean J. Gastroenterol. 2012; 59:395–400. [PubMed: 22735871]

103. Zhang L, Jamaluddin MS, Weakley SM, Yao Q, Chen C. Roles and mechanisms of microRNAs in pancreatic cancer. World J. Surg. 2011; 35:1725–1731. [PubMed: 21222121]

104. Yang D, Zhu Z, Wang W, Shen P, Wei Z, Wang C, Cai Q. Expression profiles analysis of pancreatic cancer. Eur. Rev. Med. Pharmacol. Sci. 2013; 17:311–317. [PubMed: 23426533]

105. Yuan H, Wu B, Ma S, Yang X, Yin L, Li A. Reanalysis of the gene expression profile in chronic pancreatitis via bioinformatics methods. Eur. J. Med. Res. 2014; 19 31-783X-19-31.

106. Bloomston M, Frankel WL, Petrocca F, Volinia S, Alder H, Hagan JP, Liu CG, Bhatt D, Taccioli C, Croce CM. MicroRNA expression patterns to differentiate pancreatic adenocarcinoma from normal pancreas and chronic pancreatitis. JAMA. 2007; 297:1901–1908. [PubMed: 17473300]

107. Wang B, Sun S, Liu Z. Analysis of dysregulation of immune system in pancreatic cancer based on gene expression profile. Mol. Biol. Rep. 2014; 41:4361–4367. [PubMed: 24619357]

108. He C, Jiang H, Geng S, Sheng H, Shen X, Zhang X, Zhu S, Chen X, Yang C, Gao H. Analysis of whole genomic expression profiles and screening of the key signaling pathways associated with pancreatic cancer. Int. J. Clin. Exp. Pathol. 2012; 5:537–546. [PubMed: 22949936]

109. Gorlov IP, Logothetis CJ, Fang S, Gorlova OY, Amos C. Building a statistical model for predicting cancer genes. PLoS One. 2012; 7:e49175. [PubMed: 23166609]

110. Junker, BH.; Schreiber, F. Analysis of biological networks. Hoboken, N.J: Wiley-Interscience; 2008.

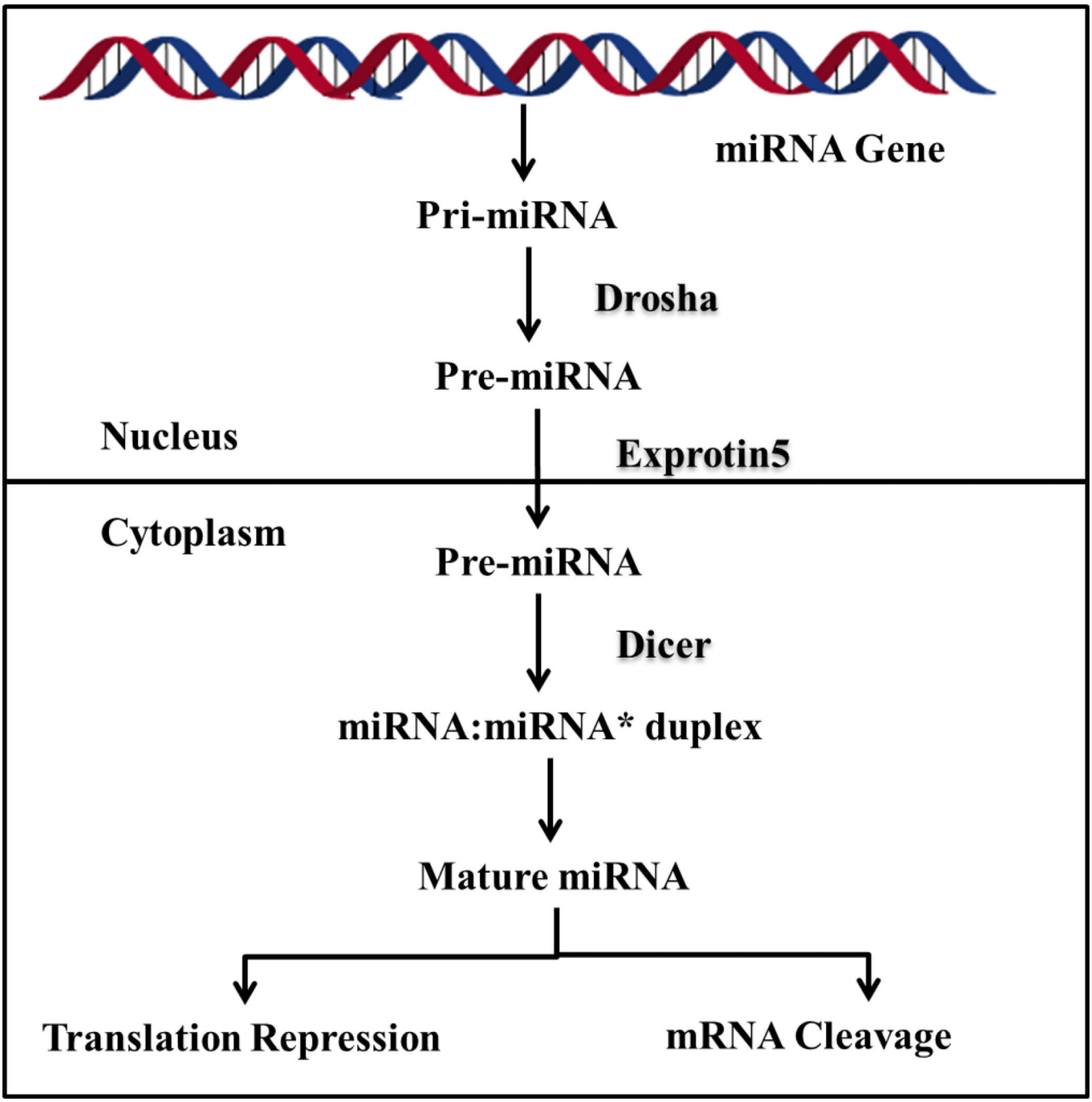

Figure 1.

miRNA biogenesis: miRNA genes are transcribed in the nucleus, and undergo subsequent processing by the endonucleases Drosha and Dicer to produce a duplex comprised of mature miRNA and its antisense strand (miRNA*). The mature miRNA strand is incorporated into the ribonucleoprotein complex (RISC), which mediates interaction with the target mRNA and mRNA silencing, either through mRNA (messenger RNA) cleavage or translational repression (Adopted from [27]).

Table 1

Text-mining of different cancer types, cancer-related miRNAs and their potential target genes with number of citations in PubMed

Type of cancer	miRNAs known to be involved in cancer type	Number of genes involved in cancer type	Number of PubMed references
Colorectal cancer	miR-143, miR-145	236	177
Pancreatic cancer	miR-21, miR-34, miR-107	512	442
Breast cancer	miR-125b, miR-145, miR-21, miR-155	1142	1577
Thyroid cancer	miR-221, miR-222, miR-146, miR-181	249	177
Acute myeloid leukemia	miR-29b, miR-191, miR-199a, miR-181a	391	439
Chronic lymphocytic Leukemia	miR-15a, miR-16-1	119	74
Basal cell carcinoma	miR-184, miR-10b, miR-98, miR-200	161	96
Melanoma	let-7b, miR-121/122, miR-137	432	345
Renal cell carcinoma	miR-141, mir-200c	209	164
Bladder cancer	miR-126, mir-182, mir-129, mir-143, miR-127, mir-125b	190	147
Prostate cancer	mir-34a, miR-21,	692	711
Endometrial cancer	miR-129-2, miR-194, miR-204	166	119
Lung cancer	Let-7, miR-17-92	691	652

Table 2

Comparison of computational tools for miRNA prediction and their main characteristics

	Tools	Strengths	Website	Weakness
Sequence or structure -based	miRscan	Used both sequence and structure based features to identify miRNA genes. It also incorporates the sequence conservation in *C. briggsae*. The predicted miRNA genes were validated through Northern blots.	Only conservation across two species was used to identify the miRNA genes.	http://genes.mit.edu/mirscan/
	miRSeeker	Identified conserved miRNA genes in *Drosophila melanogaster* that are conserved in *D. pseudoobscura* and other distant species like insects and vertebrates	Limited to Drosophila miRNA genes.	-
	miRAlign	This tool incorporates the potential of both sequence and structure alignment to predict miRNAs.	Using homolog information to identify miRNAs reduces the chances to identify non-conserved miRNAs	http://bioinfo.au.tsinghua.edu.cn/miralign/

	Tools	Strengths	Website	Weakness
Expression data based	miRDeep	Predicts miRNAs using deep sequencing data	Using the sequencing technology alone limits the search based on sequence only. Also the number of predicted miRNAs that were validated using northern blot was very few.	https://www.mdc-berlin.de/8551903/en/research/research_teams/systems_biology_of_gene_regulatory_elements/projects/miRDeep
	miRanalyzer	Uses high-throughput sequencing data as well as miRNA expression data to predict mature miRNAs or precursor miRNAs	The expression based approaches miss the miRNAs where no noted effect on gene expression happens but protein production is inhibited	http://bioinfo5.ugr.es/miRanalyzer/miRanalyzer.php
Machine learning based	miRFinder	It is a pre-miRNA prediction tool that utilizes both sequence and structure feature. It used support vector machine to predict pre-miRNAs.	Takes into account the sequence conservation across species to identify miRNAs. This may hinder the identification of species specific miRNAs	http://www.bioinformatics.org/mirfinder/
	miPred	Classifies pseudo vs real miRNA precursors using machine learning approaches. The algorithm also achieved very high accuracy for human data.	The major limitation of using the machine learning techniques is to generate a negative dataset that affects the accuracy and sensitivity of the output.	http://www.bioinf.seu.edu.cn/miRNA/

Table 3

Comparison of computational tools for miRNA target prediction and their main characteristics

Tools	Strengths	Websites	Weakness
TargetScan	Seed match and conservation	Uses conservation as a filter. This misses some of the non-conserved miRNA sites.	http://www.targetscan.org/
Diana-MicroT	Seed match, conservation, free energy, site accessibility and target-site abundance	Currently hosts miRNA target predictions for Homo sapiens, Musmusculus, Drosophila melanogaster and Caenorhabditis elegans only	http://diana.cslab.ece.ntua.gr/microT/
PicTar	Seed match, free energy and thermodynamics.	The predictions are based on very old dataset missing newly identified miRNAs.	http://pictar.mdc-berlin.de/
MirTarget2	Seed match, conservation, free energy, site accessibility and others (SVM based)	Only predicts miRNA targets for five species: human, mouse, rat, dog and chicken. It is a SVM-based approach that lacks the extensive positive and negative training datasets.	http://mirdb.org/miRDB/
MiRanda	Seed match, conservation, and free energy	It needs to be downloaded and doesn't provide any scores associated with the predictions.	http://www.microrna.org/microrna/home.do
RNAhybrid	Seed match, free energy, and target-site abundance	Meant for advanced users only as it requires user input and adjusting the complex settings, Not a user friendly tool for novice users.	http://bibiserv.techfak.uni-bielefeld.de/rnahybrid/

Table 4

Summary of existing computational tools and approaches in establishing the role of miRNA in cancer

***Insilico* models**	**Categories**	**Dataset used**	**Features**	**References**
Statistical /Gene Expression Models	Cancer Diagnosis	Transcriptomics data *i.e*.gene and miRNA expression profiling data	Identifies, if the behavior of the gene involved in a tumor diverges from the normal.	[74, 75, 80]
	Cancer types and Subtypes	Transcriptomics data *i.e*.gene and miRNA expression profiling data	Classifies the cancer types and its sub types based on the differential gene expression in a given condition	[71, 73–75,81–84]
	Cancer gene predictions	Transcriptomics data *i.e*.gene and miRNA expression profiling data	Identifies potential set of genes that might be involved in various tumors based on their change in expression pattern in tumor condition as compared to normal.	[74, 75]
Network-based Models	Functional enrichment analysis	List of differentially expressed genes in tumor condition	Gene set enrichment analysis to identify enriched pathways in cancer	[85–87]
	Gene co-expression networks	Integrated gene coexpression, transcriptional and posttranscriptional regulation network.	Cancer diagnosis and prognosis. Also helps to associate various regulatory elements in cancer	[88–91]
Biochemical Reaction Modeling	Modeling metabolic pathways	Metabolic reactions data and enzyme kinetics data	Simulates the behavior of genes or a system in cancer conditionsthus helps to understand the cancer biology at systems level	[5, 92]

Table 5

Deregulated miRNAs in Pancreatic Cancer—up and down-regulated miRNA in pancreatic cancer and their respective reference are provided below. Additionally, the table also includes the recently identified deregulated miRNAs in pancreatic cancer studies.

miRNAs	Expression Profile	References
miR-132	Up	Park *et al* 2011 [96]
miR-18a	Up	Morimura *et al* 2011[97]
miR-185	Up	Liu *et al* 2012 [80]
miR-191	Up	Liu *et al* 2012 [80]
miR-20a	Up	Liu *et al* 2012 [80]
miR-211	Up	Giovannetti*et al* 2012 [98]
miR-25	Up	Liu *et al* 2012 [80]
miR-34b	Up	Liu *et al* 2013 [99]
miR-141	Down	Zhao *et al* 2013[100]