Advances in Experimental Medicine and Biology 1385

Ulf Schmitz
Olaf Wolkenhauer
Julio Vera-González *Editors*

Systems Biology of MicroRNAs in Cancer

Springer

Advances in Experimental Medicine and Biology

Volume 1385

Advances in Experimental Medicine and Biology provides a platform for scientific contributions in the main disciplines of the biomedicine and the life sciences. This series publishes thematic volumes on contemporary research in the areas of microbiology, immunology, neurosciences, biochemistry, biomedical engineering, genetics, physiology, and cancer research. Covering emerging topics and techniques in basic and clinical science, it brings together clinicians and researchers from various fields.

Advances in Experimental Medicine and Biology has been publishing exceptional works in the field for over 40 years, and is indexed in SCOPUS, Medline (PubMed), EMBASE, BIOSIS, Reaxys, EMBiology, the Chemical Abstracts Service (CAS), and Pathway Studio.

2021 Impact Factor: 3.650 (no longer indexed in SCIE as of 2022)

Ulf Schmitz • Olaf Wolkenhauer •
Julio Vera-González
Editors

Systems Biology of MicroRNAs in Cancer

Springer

Editors
Ulf Schmitz
Department of Molecular and Cell Biology, Centre for Tropical Bioinformatics and Molecular Biology
James Cook University
Townsville, QLD, Australia

Julio Vera-González
Department of Dermatology
Friedrich-Alexander-Universität Erlangen-Nürnberg and University Hospital Erlangen
Erlangen, Bayern, Germany

Olaf Wolkenhauer
Department of Systems Biology and Bioinformatics
University of Rostock
Rostock, Germany

ISSN 0065-2598 ISSN 2214-8019 (electronic)
Advances in Experimental Medicine and Biology
ISBN 978-3-031-08355-6 ISBN 978-3-031-08356-3 (eBook)
https://doi.org/10.1007/978-3-031-08356-3

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

About the Editors

Ulf Schmitz's research focus is on RNA biology and systems medicine. He integrates in silico, in vitro, and in vivo approaches in advanced interdisciplinary research. His team develops integrative systems medicine workflows along with tools and databases to identify and study interactions between RNAs and other molecules, mechanisms of gene regulation, and gene regulatory networks.

With the goal of achieving a deeper understanding of processes involved in disease emergence and progression, his projects focus on different cancer entities and possible avenues for therapeutic interventions and optimized treatment schedules.

Olaf Wolkenhauer's research combines data-driven modelling with model-driven experimentation, using a wide range of approaches, including machine learning, statistics, systems theory, and stochastic processes. He has extensive experience in developing systems theoretic concepts and developing algorithms, workflows, standards, and software tools that support the analysis data with applications in biomedicine. His specific interest is in understanding the functioning of cells within a tissue and how such whole-part relationships can be used to identify principles underlying tissue (mal)functioning.

Julio Vera-González is a physicist working in medical systems biology since 2005. His expertise is in mathematical modelling, bioinformatics, and network biology. He applies multi-criteria decision algorithms to patient classification and therapy assessment. His primary interest is the role of (non-coding) RNA regulation in the interplay between cancer and the immune system, as well as its impact on (immuno)therapy.

Chapter 1
The Role of MicroRNAs in Cancer Biology and Therapy from a Systems Biology Perspective

Xin Lai, Ulf Schmitz, and Julio Vera

Abstract Since the discovery of microRNAs (miRNAs) in *Caenorhabditis elegans*, our understanding of their cellular function has progressed continuously. Today, we have a good understanding of miRNA-mediated gene regulation, miRNA-mediated cross talk between genes including competing endogenous RNAs, and miRNA-mediated signaling transduction both in normal human physiology and in diseases.

Besides, these noncoding RNAs have shown their value for clinical applications, especially in an oncological context. They can be used as reliable biomarkers for cancer diagnosis and prognosis and attract increasing attention as potential therapeutic targets. Many achievements made in the miRNA field are based on joint efforts from computational and molecular biologists. Systems biology approaches, which integrate computational and experimental methods, have played a fundamental role in uncovering the cellular functions of miRNAs.

In this chapter, we review and discuss the role of miRNAs in oncology from a system biology perspective. We first describe biological facts about miRNA genetics and function. Next, we discuss the role of miRNAs in cancer progression and review the application of miRNAs in cancer diagnostics and therapy. Finally, we elaborate on the role that miRNAs play in cancer gene regulatory networks. Taken together, we emphasize the importance of systems biology approaches in our continued efforts to study miRNA cancer regulation.

Author contribution: Conceptualization: XL, US, JV; Visualization: XL; Writing-original draft: XL, JV; Writing-review and editing: XL, US, JV.

X. Lai (✉) · J. Vera (✉)
Laboratory of Systems Tumor Immunology, Department of Dermatology, Friedrich-Alexander-Universität Erlangen-Nürnberg (FAU) and Universitätsklinikum Erlangen, Erlangen, Germany

Deutsches Zentrum Immuntherapie and Comprehensive Cancer Center Erlangen-EMN, Erlangen, Germany
e-mail: xin.lai@uk-erlangen.de; julio.vera-gonzalez@uk-erlangen.de

U. Schmitz
Department of Molecular and Cell Biology, College of Public Health, Medical and Veterinary Sciences, James Cook University, Townsville, QLD, Australia

U. Schmitz et al. (eds.), *Systems Biology of MicroRNAs in Cancer*, Advances in Experimental Medicine and Biology 1385,
https://doi.org/10.1007/978-3-031-08356-3_1

Keywords mRNA destabilization · Translation repression · Mathematical modelling · Network biology · Bioinformatics · Posttranscriptional gene regulation · Oncomir · Tumor suppressor miRNA · Metastamir · Circulating miRNA

1.1 Biological Facts About miRNA Biogenesis and Function

MicroRNAs (miRNAs) are a class of small endogenous noncoding RNAs with a length of around 22 nucleotides (nt). miRNAs are evolutionarily conserved regulatory molecules that, in most cases, modulate the stability and/or translation of target mRNAs through direct binding to the target's 3′ UTR (Bartel 2004). miRNAs were first found to be pivotal for *Caenorhabditis elegans* development (Vella et al. 2004) and it was soon demonstrated that they play a key role in gene expression regulation in both animals and plants. More recently, there is mounting evidence suggesting that miRNAs and other similar noncoding RNAs are also important in viral and bacterial gene regulation, as well as in the microbe-mediated host gene regulation. Taken together, miRNAs are ubiquitous posttranscriptional regulators of gene expression and important in normal cell physiology and function (Cardin and Borchert 2017).

To date over 2500 miRNA sequences have been identified in the human genome and registered in the miRBase database (Kozomara et al. 2019). These miRNAs are estimated to regulate more than half of all protein-coding genes (Friedman et al. 2009). This indicates their pervasive roles in the regulation of cellular processes, like proliferation, differentiation, and apoptosis. In addition to exerting critical functions during normal development and cellular homeostasis, miRNA dysregulation has been found in many human diseases, like cancer (Hwang and Mendell 2006). Thus, understanding the function of miRNAs in gene regulation is crucial for unraveling mechanisms underlying human pathogenesis and improving therapeutic approaches in human diseases.

1.1.1 miRNA Biogenesis

The miRNA biogenesis pathway is a complex process composed of multiple steps (Fig. 1.1) (Berezikov 2011; Filipowicz et al. 2008; Krol et al. 2010). At first, long primary transcripts known as primary miRNAs (pri-miRNAs) are transcribed from miRNA genes by RNA polymerase II (Pol II). Pri-miRNA molecules have a 5′-terminal 7-methylguanosine (m^7G) cap, which is extended by a hairpin structure with a terminal loop and a ~32 nt long imperfectly base-paired stem and end with a 3′ poly(A) tail. Depending on the features of miRNA genes, pri-miRNAs can contain single or multiple miRNA pairs that form hairpin structures. Next, with the help of the complex that includes Drosha and its binding partner DGCR8, pri-miRNAs are

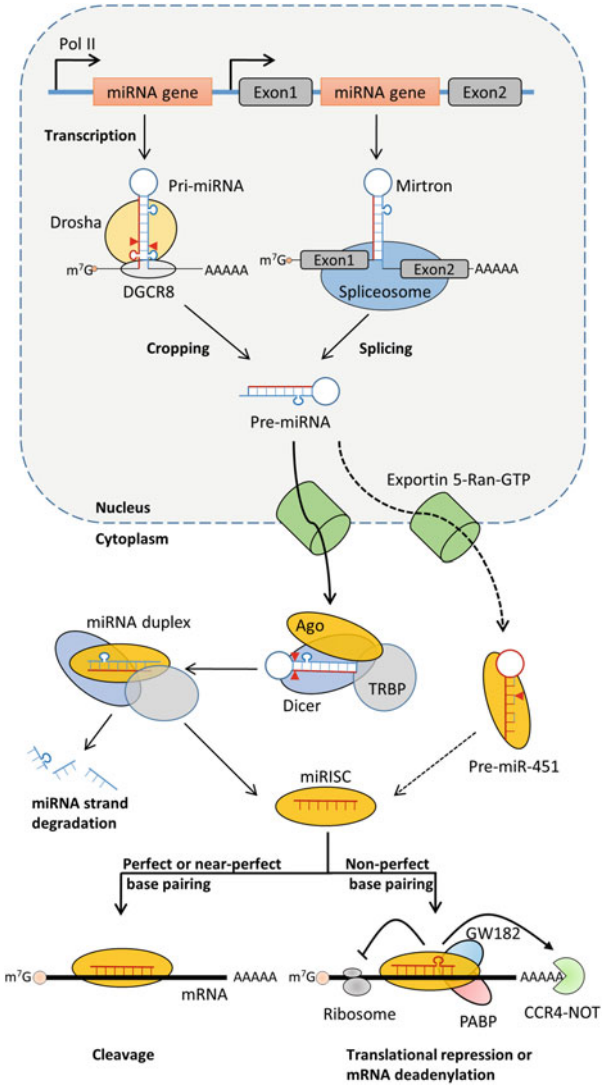


Fig. 1.1 miRNA biogenesis. A miRNA can be processed either from a pri-miRNA or a mirtron. The pri-miRNA, containing a 5′ terminal m^7G cap and a 3′ poly(A) tail (AAAAA), is transcribed from miRNA genes by Pol II and is subsequently cleaved (red arrowheads) by Drosha with the cofactor DGCR8 and becomes a pre-miRNA. The mirtron situated between two exons is spliced and becomes a pre-miRNA without the requirement of Drosha–DGCR8 complex. The pre-miRNA is transported from the nucleus to the cytoplasm by exportin 5 with Ran-GTP. In the cytoplasm, most pre-miRNAs are processed into double-stranded miRNA duplexes with the help of Dicer and TRBP. One strand of the duplexes is loaded into AGO, whereas the other strand is degraded. When a miRNA is perfectly or near-perfectly pairing to its target mRNA, it can result in the cleavage of the mRNA. Otherwise, non-perfect base pairing between a miRNA and its target mRNA leads to translation repression or target mRNA deadenylation. Both processes are implemented through the interaction of miRISC with GW182 and PABP

processed into precursor miRNAs (pre-miRNAs), which are ~70 nt long hairpin structures with a characteristic 2 nt 3′ overhang. Then, through the recognition of the 2 nt overhangs, exportin 5 in conjunction with the cofactor Ran-GTP exports pre-miRNAs from the nucleus into the cytoplasm. After that, cytoplasmic processing by another complex, which is composed of Dicer, an Argonaute protein (AGO), and a TAR RNA binding protein (TRBP), cleaves the pre-miRNA into a ~22 nt double-stranded miRNA duplex (also known as mature miRNAs). Finally, one strand of the miRNA duplex known as the active strand is loaded into the AGO-containing miRNA-induced silencing complex (miRISC) that will bind to miRNA-specific target mRNAs for subsequent cleavage or translation repression. The complementary strand of the miRNA duplex, known as the passenger strand, will be degraded.

In addition to the canonical miRNA biogenesis pathway described above, mature and functional miRNAs can also be produced via alternative pathways. These pathways can be classified into Drosha- and Dicer-independent pathways (Fig. 1.1) (Miyoshi et al. 2010). In the Drosha-independent pathway a class of miRNA genes, which originates from pre-miRNA-sized short introns (termed as mirtrons), can be directly processed into pre-miRNA hairpins without the participation of Drosha. These pre-miRNAs are further cleaved by Dicer in the cytoplasm to produce mature miRNAs (Ruby et al. 2007). In the Dicer-independent pathway, following normal nuclear processing, the pre-miRNA is not cleaved into a miRNA duplex by Dicer but instead by the AGO catalytic center. For example, miR-451 is produced through an AGO-dependent maturation pathway (Miyoshi et al. 2010).

1.1.2 miRNA Function

After the maturation of miRNAs, in most cases the active strands act as guides and direct miRISCs to bind to the 3′ UTR of target mRNAs, resulting in the repression of target genes at the posttranscriptional level (Fig. 1.2). Some miRNAs can exert a repressive function on target genes even when their binding sites are located in the 5′ UTR or the coding regions of target mRNAs (Lytle et al. 2007). In addition, a few miRNAs can bind to the 5′ UTR of their target mRNA and enhance its translation (Ørom et al. 2008). The mechanism by which target mRNAs are regulated is determined by the degree of complementarity between miRNAs and their target mRNAs. When a miRNA perfectly or near-perfectly pairs with its target mRNA, mostly occurring in plants, target mRNA cleavage is triggered. Imperfect base pairing between a miRNA and its target, predominating in animals, leads to translation repression or destabilization of the target mRNA (Bartel 2004). Based on experimental evidence and bioinformatics analyses in animals, several miRNA seed binding motifs have been identified including 8-mer, 7-mer, and 6-mer seed binding (Bartel 2009). These miRNA binding motifs are defined by the number of continuous base pairings in the seed region of miRNAs; for example, 7-mer means that in the seed region of a miRNA there are seven continuous base pairings between the miRNA and its target mRNA (Bartel 2009). The repression efficiency exerted via

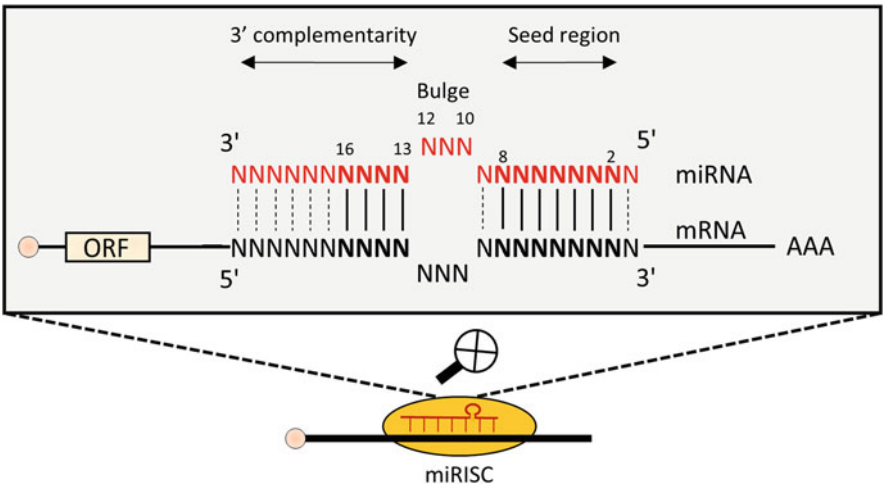


Fig. 1.2 Illustration of base pairing between miRNAs and target mRNAs. In the seed region, continuous Watson-Crick pairing (vertical solid lines) is crucial for the efficient duplex formation and miRNA-mediated repression. When a mismatch (vertical dashed lines) or a bulge appears in the seed region, Watson-Crick pairing centering on miRNA nucleotides 13–16 of the 3′ complementarity region can compensate and thereby construct a functional miRNA binding site

these binding motifs can be further enhanced by additional base pairing between the 3′ complementarity region of the miRNA and its target (Fig. 1.2) (Filipowicz et al. 2008). The following subsequences can influence target regulation efficiency: (1) The seed region (miRNA nucleotides 2–8). A continuous seed region base pairing (miRNA nucleotides 2–8) is crucial for assuring effective target repression. If there are G-U pairs (guanine-uracil) or mismatches in this region, the target repression will be greatly affected. However, the appearance of an A (adenine) at position 1 of the miRNA and an A or U appearing at position 9 can improve the repressive efficiency, although they are not required to base pair with the target mRNA. (2) The central region (miRNA nucleotides 10–12). In this region, bulges or mismatches must be present. (3) The complementary region (miRNA nucleotides 13 to last). The base pairing between the miRNA and target mRNA is typically quite loose in this region. However, good complementarity, particularly for miRNA nucleotides 13–16, becomes important when mismatches or bulges appear in the seed region.

In addition to the binding motifs, other factors can also affect miRNA repression efficiency. For example, multiple miRNA binding sites in close proximity in the 3′ UTR of a single mRNA can enhance the repression of the target (Doench and Sharp 2004; Saetrom et al. 2007). RNA-binding proteins (RBPs), which can interact with miRISCs on the 3′ UTR of target mRNAs, can either facilitate or counteract miRNA-mediated repression (Krol et al. 2010).

In mammalian cells, most miRNA–target interactions are based on imperfect base pairing, which can result in two main mechanisms by which miRNAs reduce protein

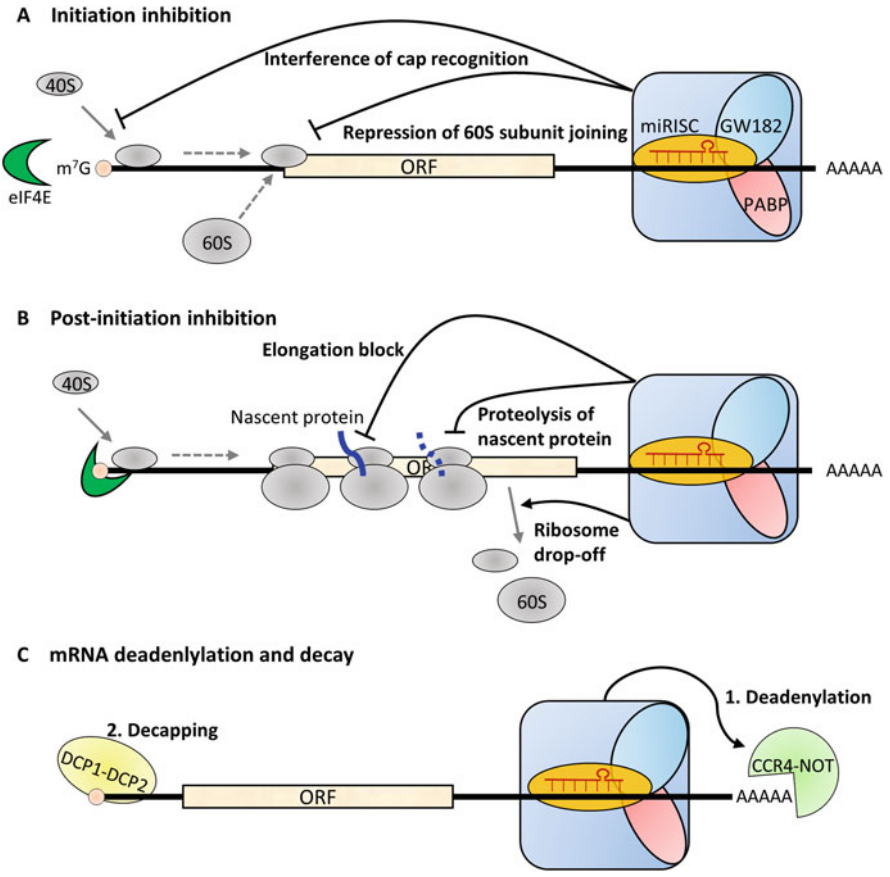


Fig. 1.3 miRNA-mediated translation repression mechanisms. With the help of GW182 and PABP, miRISCs can repress translation at the initiation and post-initiation stage, or induce the deadenylation and decay of target mRNAs. (**a**) At the initial stage, binding of the miRISC complexed with GW182 and PABP to the target mRNA can repress translation by either interfering with the cap recognition or by repressing the 60S subunit joining. (**b**) The miRISC can inhibit translation at the post-initiation step by blocking translation elongation, causing ribosome drop-off or proteolytic cleavage of the nascent polypeptides. (**c**) Deadenylation of the target mRNA is facilitated by the interaction of the miRISC with CCR4-NOT. Subsequently, the decay of the target mRNA happens after the removal of the 5′-terminal m7G cap by the decapping DCP1-DCP2 complex

production. The two mechanisms are translation repression and destabilization of the target mRNAs. More particularly, miRNAs can inhibit the translation of target mRNAs by affecting the initiation or post-initiation stage of mRNA translation (Fabian et al. 2010). At the initiation stage, the miRISC can inhibit translation by interfering with eIF4E-cap recognition and recruitment of 40S small ribosomal subunit or by antagonizing 60S subunit joining and preventing the formation of 80S ribosomal complex (Fig. 1.3a). At the post-initiation stage, the miRISC can

inhibit translation by blocking ribosome elongation, inducing ribosome drop-off, or facilitating proteolysis of the nascent polypeptides (Fig. 1.3b). For miRNA-mediated mRNA degradation, with the participation of GW182 and PABP the miRISC induces deadenylation of the poly (A) tail by interacting with the CCR4-NOT deadenylase complex. Then, the 5′ terminal m^7G cap is removed by the DCP1-DCP2 decapping complex, resulting in the degradation of the target mRNA (Fig. 1.3c).

What we described above can be considered the "standard" mechanism of gene regulation by miRNAs. However, recent experimental studies indicate that miRNAs can translocate into the nucleus, interact with gene promoters and activate expression for given target genes (Xiao et al. 2017). Xiao et al. found that nuclear miR-24-1 can activate gene transcription in a mechanism that involves interaction with enhancers and enhancer RNAs. This indicates that in the future, once we elucidate the mechanisms of interaction between miRNAs, RNA binding proteins, and long noncoding RNAs (lncRNAs), we may see more cases of alternative miRNA target regulation.

1.2 The Role of miRNAs in Cancer Progression, Diagnosis, and Therapy

In the previous section, we discussed the genetics and molecular mechanisms associated with miRNA biogenesis and function. The large amounts of experimental evidence collected in the last 20 years show that miRNAs participate in the regulation and fine-tuning of crucial processes that drive cellular phenotypes and functions during cell development and repairing and in tissue homeostasis (Gangaraju and Lin 2009; Piotto et al. 2018). Since cancer cells hijack cell differentiation programs to regain phenotypes that foster their progression, it is not surprising that many miRNAs are associated with the pathogenesis and progression of cancer. In recent years, researchers have investigated the phenomenon to search for new, accurate diagnostic tools based on miRNA expression profiling in cancer patients. Furthermore, there are miRNA-based therapies under development, which promote more targeted and personalized cancer therapies. In the following paragraphs, we discuss the molecular mechanisms linking miRNAs to cancer pathophysiology and the use of miRNAs in cancer diagnostics and therapy.

1.2.1 Genome-Level Alterations in miRNAs

As indicated before, miRNAs play a key role in shaping and fine-tuning the gene regulatory circuits controlling tissue development and cell differentiation. Programs controlling these phenotypes are hijacked by cancer cells allowing them to become invasive, metastatic, and therapy resistant. Hence, one can find mechanisms by

which miRNA expression and function are distorted in cancer cells similar to those that cause dysregulation of protein-coding genes.

Alterations in miRNA gene copy number and gene location Comprehensive genome analyses using computational and experimental approaches have identified a large number of miRNA genes that are located in fragile and unstable chromosomal regions linked to cancer. One can find miRNA genes within or close to cancer-associated amplified, deleted, or translocated genes, but also close to chromosomal breakpoints. For example, Soh and co-workers analyzed genomic data from more than 2000 tumor samples of The Cancer Genome Atlas (TCGA) cohort representing seven prevalent cancer types and found that up to 85% of miRNA genes are located in cancer type-specific genomic regions enriched in somatic copy number alterations (SCNAs) (Soh et al. 2018). Czubak et al. investigated SCNAs of 14 miRNA genes commonly deregulated in cancer and found that most of them displayed copy number alterations in lung cancer (Czubak et al. 2015).

Alterations in miRNA biogenesis and transcriptional regulation Global miRNA depletion caused by epigenetic or genetic alterations in miRNA biogenesis components is oncogenic (Lin and Gregory 2015). For example, DICER1 seems to operate as a haploinsufficient tumor suppressor. Kumar et al. showed that deleting a single copy of DICER1 in tumors from animal models reduced survival (Kumar et al. 2009). Others made similar observations in cell lines and mouse models of several aggressive cancers (Poma et al. 2019). Reduced expression of DICER1 has been associated with various cancers while it is not a general pattern and therefore, it is difficult to generalize the role of DICER repression or silencing in cancer (Foulkes et al. 2014). This can also be found for other proteins involved in the miRNA biogenesis pathway, which are considered either tumor suppressors or oncogenes for different tumors (Hata and Kashima 2016). For example, DROSHA and DGCR8 both critical components of the miRNA biogenesis pathway, bear recurrent mutations in some cancers (Walz et al. 2015).

The expression of many miRNAs is also controlled by transcription factors (TFs) and cofactors, some of which are commonly deregulated in cancer. Consequently, miRNAs regulated by those TFs experience cancer type-specific alterations in their expression patterns. The literature contains many studies in which the expression of miRNAs is controlled by TFs that are deregulated or mutated in cancer. To mention a few, there is the tumor suppressor p53 which regulates the expression of the miR-34 family, the translational repressor ZEB1 regulating miR-200 family expression, tumor suppressor p73 regulating miR-205 expression, and the proto-oncogene c-Myc regulating the miR-17 ~92 cluster. In all these cases, evidence shows that cancer-associated deregulation of these TFs induces abnormal expression of the target miRNAs, which in turn promote post-transcriptional repression of genes linked to key cancer phenotypes, such as cell proliferation, (anti)apoptosis, or migration. In one particular case, the TF p73 promotes the expression of miR-205, a miRNA involved in the repression of several anti-apoptotic members of the BCL2 family. Deregulation of p73 expression in cancer downregulates miR-205, which in turn induces an increased level of anti-apoptotic BCL2 (Vera et al. 2013; Alla et al.

2012). In such a scenario, the cells initiate apoptosis following DNA damage and become resistant to genotoxic drugs.

There are several interesting points here. Firstly, deregulation of these TF can be induced in different ways, e.g., via somatic mutations, overexpression, or alternative splicing. The latter, for example, causes the expression of the anti-apoptotic DNp73 splice isoform instead of wild-type p73. Secondly, some TFs such as p53 or p73 can upregulate the expression of miRNAs, and others can repress miRNA expression like ZEB1 represses the miR-200 family. Finally, in many cases, miRNAs and TFs are part of feedback and feedforward loops becoming deregulated in cancer. For example, miR-205 establishes a negative feedback loop with E2F1 and p73 leading to therapy resistance in malignant melanoma (Vera et al. 2013). Taken together, the interplay between miRNAs and TFs is a complex multifactorial mechanism, whose features and consequences will be discussed in more detail in Sect. 1.3.

Modification and deregulation of miRNA–mRNA interactions Not surprisingly, the binding between miRNAs and their target mRNAs can also be distorted in cancer. Binding can be altered through genetic changes in target genes such as somatic point mutations or translocations. A point mutation within the canonical seed-matching sequence of the mRNA 3′ UTR could create a novel miRNA target site but it could also impair an existing target site (Moszyńska et al. 2017). For example, a point mutation located in the 3′ UTR of the p53 inhibitor MDM4 (rs4245739 SNP, A > C) is associated with an increased risk of prostate cancer. Bioinformatics analysis indicated that this SNP resides within a predicted binding site for miR-191-5p, miR-887, and miR-3669. Stegeman and coworkers investigated these predictions utilizing gene assays and demonstrated that miR-191-5p and miR-887 have a specific affinity for the rs4245739 SNP C-allele in prostate cancer. When targeting MDM4 with miR-191-5p or miR-887 in prostate cancer cell lines they observed decreased cell viability (Stegeman et al. 2015). There are other mechanisms altering the molecular structure of miRNA–mRNA binding sites, including chromosomal translocations that eliminate given miRNA binding sites from the 3′ UTR of their mRNA targets (Hirano et al. 2019), or alternative polyadenylation, which can shorten or lengthen a gene's 3′ UTR and thereby erase or add miRNA binding sites, respectively (Mao et al. 2020).

1.2.2 Oncogenic and Tumor-Suppressive miRNAs

In the previous sections, we have discussed miRNA biogenesis and target repression mechanisms and we have elucidated mechanisms by which miRNAs can become deregulated in cancer. Now, the actual role that given miRNAs play in cancer will depend largely on two factors: (1) whether they are up-or downregulated, and (2) the function of their gene targets.

Oncogenic miRNAs, also known as oncomirs, are miRNAs that repress genes with a known role as a tumor suppressor. For example, miR-125b has a binding site

in the 3′ UTR of the tumor suppressor TP53. It has been found that miR-125b overexpression reduces the endogenous levels of TP53 and hampers the activation of apoptosis in several cancers (Le et al. 2009). Of note are also miRNAs that repress genes that inhibit or downregulate oncogenic pathways. This is the case for miR-663, which represses the expression of CDKN1A, a gene that encodes the cell cycle regulator p21. The protein can induce cell cycle arrest, however, miR-663-mediated suppression of p21 promotes cancer cell proliferation and tumor progression in nasopharyngeal carcinoma and other cancers (Yi et al. 2012).

Other miRNAs described as tumor suppressors target oncogenes. A well-known example is the role of the let-7 family in melanoma. These miRNAs suppress melanoma proliferation and metastasis by targeting a range of genes including ITGB3, an integrin linked to the acquisition of invasiveness (Müller and Bosserhoff 2008). Such kinds of miRNAs can also target inhibitors of tumor suppressors. A well-known case is miR-34a which represses SIRT1 expression. SIRT1 is an oncogene that would normally repress TP53 activation. It has been shown that miR-34a-mediated repression of SIRT1 increases TP53 acetylation and hence the expression of TP53 target genes (such as CDKN1A and PUMA) that regulate cell cycle and apoptosis (Yamakuchi et al. 2008). More recently, it has been shown that miRNAs also play a role in regulating cancer therapy efficiency and resistance to chemotherapy. For example, Alla and coworkers found that DNp73-dependent downregulation of miR-205 induces drug resistance by upregulating anti-apoptotic BCL2 and ABC transporters (Alla et al. 2012).

Taken together, miRNAs play a crucial role in cancer through the inhibition of tumor suppressors or oncogenes. Interestingly, since miRNAs can have multiple targets, some miRNAs play contradictory roles in different tumor entities or even within same cancer. For example, miR-146a can promote melanoma cell growth by targeting NUMB, a repressor of the NOTCH signaling pathway (Forloni et al. 2014), but can also suppress metastasis formation by downregulating the expression of ITGAV and ROCK1 (Raimo et al. 2016).

1.2.3 miRNAs in Cancer Diagnostics and Therapy

In the last decade, a lot of work has been carried out to find means to use miRNAs for primary or co-adjuvant therapies but also to identify biomarkers to predict disease outcomes or resistance to therapy.

miRNAs as diagnostic signatures Due to the lack of sufficient specificity and sensitivity of classical tumor biomarkers, researchers have been looking for alternative candidates for cancer diagnosis. A good alternative should be minimally invasive and cost-effective. Profiling of circulating miRNAs from liquid biopsies has been found to be a good means to identify tumor-derived molecules secreted into the bloodstream. These miRNAs are good candidates for biomarkers because they are chemically stable and resistant to RNase activity (Glinge et al. 2017) and are thus a

valuable source for the diagnosis and stratification of cancer subtypes (Quackenbush et al. 2014). A paradigmatic example is a work by van Laar et al. (Van Laar et al. 2018). The authors utilized the Nanostring nCounter system to perform extensive profiling and quantification of miRNAs in plasma samples from melanoma patients and healthy controls. After analyzing the data, they identified 38 miRNAs that were differentially expressed between melanoma and healthy plasma samples. Interestingly, most of these miRNAs regulate protein-coding genes linked to angiogenesis, metastasis, or therapy resistance, including miR-34a and miR-205 that were discussed in the previous section. To test the prediction accuracy of this miRNA signature, they trained a machine-learning model and validated it using additional independently published datasets. The results indicated a high classification performance (with the receiver operator characteristic curve value of 0.94). Blood profiling of miRNAs is theoretically applicable for any tumor type, and in some specific tumor entities, it is also possible to profile miRNAs in urine (Sapre et al. 2016) or feces (Duran-Sanchon et al. 2020) for diagnostic purposes.

miRNA-based therapy The increasing knowledge about the roles of miRNAs in the pathogenesis, progression, and dissemination of tumors makes them attractive targets for cancer therapeutic approaches. As indicated above, miRNAs can contribute to cancer progression by acting as either oncogenes or tumor suppressors. This informs the design of miRNA-based therapies of which there are two different approaches:

1. If the aim is the inhibition of oncogmirs, one can utilize RNA antagonists, such as antisense oligonucleotides, antagomirs, or miRNA sponges. miRNA antagonists designed with sequences complementary to oncomirs prevent them from binding to AGO, thereby avoiding the inhibition of their tumor suppressor gene targets. For example, miR-146a negatively regulates immune activation by repressing STAT1 and the STAT1-dependent secretion of interferon-γ. miR-146a levels have been found to be increased in the microenvironment of aggressive melanoma. Mastroianni and coworkers combined an anti-mir for miR-146a and anti-PD1 therapy in a melanoma mouse model and found improved survival when compared with both isotype-control or anti-PD-1 treatment alone (Mastroianni et al. 2019).
2. If the treatment aims to replace depleted tumor suppressor miRNAs, one can employ miRNA mimics, like miRNA expression vectors and synthetic double-strand miRNAs. miRNA mimics can restore the diminished or lost function of tumor suppressor miRNAs whose downregulation results in the activation of oncogenes or pathways. As discussed earlier, miR-205 is a tumor suppressor and is downregulated in several aggressive tumors including melanoma. Noguchi and coworkers developed a chemically modified synthetic miRNA-205 with the ability to inhibit melanoma growth and progression, which they could demonstrate both in vitro and in vivo. They found that the synthetic miRNA can downregulate the expression of known miR-205 targets such as E2F1 and VEGF, and repress the anti-apoptotic gene BCL2.

As discussed in the previous section, a miRNA has a multitude of targets, and therefore we have to consider the potential off-target effects caused by any miRNA therapies (Ishida and Selaru 2013; Lai et al. 2019). Furthermore, similar to any other RNA-based therapies, a key challenge for miRNA therapeutics is the development of efficient delivery systems that facilitate a safe and effective application, which is a field under active investigation (Baumann and Winkler 2014; Winkle et al. 2021). However, the wide use of mRNA vaccines for protecting us from SARS-CoV-2 has laid the foundation for inventing successful carriers for therapeutic RNAs (Buschmann et al. 2021).

1.3 miRNAs in Cancer Gene Regulatory Networks

We have discussed the molecular mechanisms for miRNA (de)regulation and function in cancer. We can apprehend the complexity we face when trying to understand and exploit the therapeutic role of miRNAs in cancer. In the following, we introduce and discuss some miRNA-related phenomena that complicate miRNA-gene regulatory circuits in cancer.

1.3.1 miRNA Clusters: Groups of Similarly Regulated miRNAs

A miRNA cluster is a group of miRNA genes residing in close proximity in the genome (Lai and Vera 2013a). To consider a group of miRNAs as a cluster they have to (a) be transcribed in the same orientation, (b) they are not separated by another transcriptional unit, or (c) a miRNA on the opposite strand. There are approximately 160 miRNA clusters in the human genome. Most of the miRNA clusters are composed of two or three individual miRNAs, though larger ones are possible and often contain miRNA sets with important regulatory functions. For example, the miR-17-92 cluster contains six miRNA genes that reside in an intron of a 7-kb long noncoding RNA known as the MiR-17-92a-1 Cluster Host Gene or *MIR17HG* for short (Fig. 1.5a). Members of this cluster are the miRNAs miR-17, miR-18a, miR-19a, miR-20a, miR-19b-1, and miR-92-1. The cluster can be activated by the TF c-Myc, a well-known oncogene that is often constitutively expressed in cancer and is primarily, but not only, linked to abnormal cell proliferation. Thus, the cluster plays an important role in many cancers. One of its most relevant targets is the tumor suppressor and cell cycle regulator PTEN, which is commonly suppressed in cancer (Fuziwara and Kimura 2015).

1.3.2 Target Hubs: Genes Regulated by Many miRNAs

miRNA target hubs are genes that are regulated by 10 or more miRNAs, that is, they contain *bona fide* binding sites in their 3′ UTR for ten or more different miRNAs (Lai and Vera 2013b). Shalgi and coworkers utilized a computational approach to detect miRNA target hubs in the human genome and found 470 genes potentially regulated by at least 15 different miRNAs (Shalgi et al. 2007). Since the number of known miRNAs has almost doubled since 2007, one can expect to find many more miRNA target hubs. Since one miRNA alone often induces only mild repression of its gene target, multiple miRNAs with the same target can induce stronger repression when acting in a concerted manner. The first example of a miRNA target hub detected and experimentally investigated is the cell cycle regulator CDKN1A (Fig. 1.5b). This gene can induce cell cycle arrest under normal conditions and in response to DNA damage and is therefore considered a tumor suppressor gene. Interestingly, through bioinformatics analysis researchers detected several dozen binding sites for different miRNAs in the CDKN1A 3′ UTR (Lai et al. 2013). Wu and collaborators proved that at least 28 of these miRNAs can repress the gene in vitro (Wu et al. 2010). Interestingly, 8 of these 28 miRNAs originate from the chromosome 19 miRNA cluster, which is known to promote cancer proliferation and is linked to aggressive tumors (Jinesh et al. 2018). Lai and coworkers developed a mathematical model of CDKN1A regulation and simulated the concerted inhibition of CDKN1A during the cell cycle, DNA damage, cell cycle arrest, senescence, and apoptosis (Lai et al. 2013).

1.3.3 miRNA Cooperativity: Synergistic Gene Regulation by Multiple miRNAs

Two research teams have independently confirmed that miRNA pairs with binding sites that are in close proximity in a mutual target gene can show cooperative behavior. In other words, the effect of their combined repression is higher than the sum of the individual effects. Sætrom and coworkers experimentally determined the optimal distance between miRNA binding sites facilitating miRNA cooperativity. The optimal range is for the seed sites to be 13–35 nt apart. Based on this criterion, Lai and collaborators identified multiple pairs of putatively cooperating miRNAs in the 3′ UTR of the miRNA target hub CDKN1A (Lai et al. 2012a). They then validated the cooperative repression of CDKN1A exerted by miR-572 and miR-93 using a luciferase reporter system as well as immunoblotting. When they extended the computational analysis to the whole human genome, they identified thousands of putatively cooperating miRNA pairs and their mutual target genes (Schmitz et al. 2014). Since then, other groups have confirmed cooperative miRNA regulation in other genes, some of which are related to cancer (Bogusławska et al. 2018; Vandenwijngaert et al. 2018). The possibility that miRNAs act in a cooperative manner has consequences for miRNA-based therapies. Utilizing this synergistic

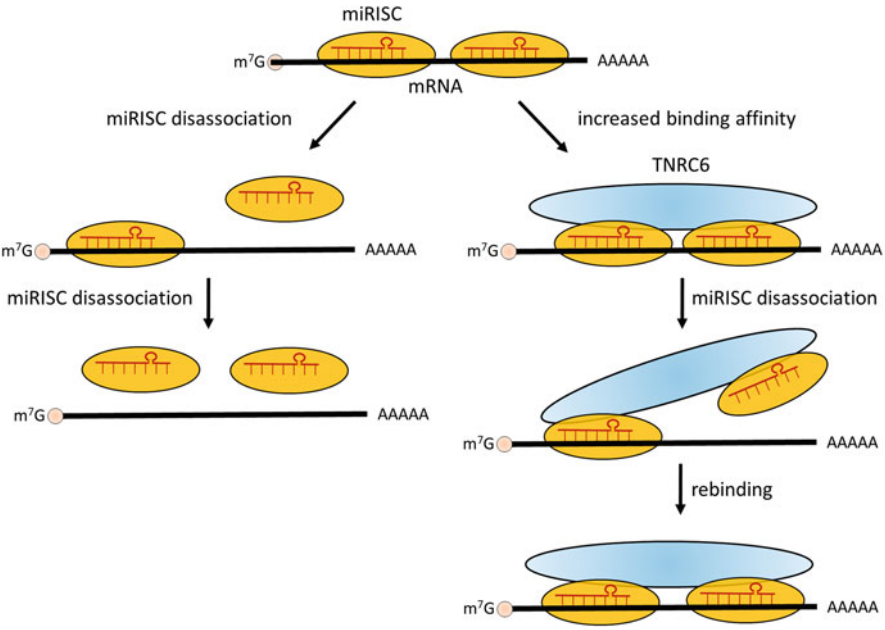


Fig. 1.4 Model of miRNA cooperativity with or without TNRC6 participation. Proximal miRNA binding sites on the target mRNA can result in cooperative gene repression by two miRNAs. TNRC6 can decrease the dissociation rate of miRISCs (formed by a miRNA and AGO) from the mRNA. Because TNRC6 simultaneously binds to two miRISCs, when one miRISC dissociates from the mRNA, TNRC6 could prevent the dissociation of the other miRISC, therefore allowing for rebinding of the disassociated miRISC

effect in a miRNA replacement therapy would reduce the overall miRNA concentration required to effectively diminish the target gene expression and thereby reduce off-target effects (Lai et al. 2019). Lai et al. explored this idea and investigated the therapeutic use of cooperative miR-205-5p and miR-342-3p in the repression of their mutual target E2F1 in the context of cancer chemoresistance (Lai et al. 2018). Their computational model-driven analysis was confirmed by in vitro functional experiments. Most recently, it was shown that the biochemical basis of miRNA cooperativity is regulated by TNRC6. The presence of the AGO-binding region in TNRC6 prevents dissociation of miRISCs from closely spaced target sites on mRNA and therefore improving their binding affinities (Briskin et al. 2020) (Fig. 1.4).

1.3.4 Network Motifs: miRNA-Enriched Feedback and Feedforward Loops

A network motif is a regulatory structure involving several genes, which recurringly appears in large biochemical networks (Alon 2007; Lai et al. 2016). One can consider miRNA target hubs as a type of network motif, but the term was coined

for feedback and feedforward loops. It is known for a long time that some central genes, especially TFs, are integrated into multiple instances of these loops and hence it is not a surprise to find a similar occurrence for miRNAs, their TFs, and their targets. Here, we discuss some examples of network motifs and their role in cancer biology.

Feedback loops In feedback loops, the regulation between molecules forming a closed loop allows state changes or self-regulation of a system. A *positive feedback loop* often induces signal amplification or sustained system (de)activation. We have found a myriad of feedback loops distorted in cancer, which integrated oncogenes and oncomirs. A well-studied case of a positive feedback loop in cancer is the one established by p53 and miR-34a with the mediation of the oncogene SIRT1 (SIRT1 $\dashv$ p53 $\rightarrow$ miR-34a $\dashv$ SIRT1). SIRT1 is overexpressed in several tumors, including melanoma, and through this circuit, it can impair the p53-mediated DNA damage and anti-proliferative response (Wilking et al. 2014; Lai et al. 2012b). miR-34a is considered a tumor suppressor and happens to be downregulated in some cancers (Zhang et al. 2019). Researchers also found an abundance of positive feedback loops involving cytokines, their signaling pathways, and downstream TFs, which play a central role in amplifying and (de)regulating the immune response in the tumor microenvironment (Jia et al. 2017). A special form of a positive feedback loop is called a toggle switch. For example, the mutual repression of a TF and its miRNA target can become a toggle switch. These motifs can display a nonlinear regulation named all-or-nothing. Specifically, the expression of one of the components represses the other in a sustained manner. The well-known and investigated case is the one established by ZEB1 and the miR-200 family (Fig. 1.5c), which plays a pivotal role in the abnormal epithelial-to-mesenchymal transition in cancer (Burk et al. 2008). A *negative feedback loop* often induces the quick cessation of signaling like the NF-kB pathway (Hoffmann et al. 2006; Inoue et al. 2016). They can also induce homeostasis and hence are employed to fine-tune signaling and gene expression and maintain levels of activity of their components against noise and fluctuation (Dublanche et al. 2006; Zhang et al. 2012). The Ras/Raf/MEK/ERK pathway is a well-known example of a pathway with multiple negative feedback loops which control cell proliferation and can get distorted in cancer (Lake et al. 2016). Not surprisingly, in recent years several research groups have found miRNAs that regulate the Ras/Raf/MEK/ERK pathway creating negative feedback loops and thereby suppressing cell growth and invasion. However, these miRNAs are often downregulated in cancer (Guo et al. 2019; Ghousein et al. 2020). There are other more sophisticated means of distortion in miRNA-mediated negative feedback loops. This is the case for the E2F1-p73/DNp73-miR-205 circuit. miR-205 represses E2F1 and is simultaneously activated by p73 a target of E2F1, therefore forming a negative feedback loop (Fig. 1.5d) (Vera et al. 2013). However, in some cancers, there is a shift toward an alternative splice isoform of p73 named DNp73, which represses miR-205 and amplifies E2F1 expression. This is often observed in aggressive tumors, such as malignant melanoma and lung adenocarcinoma, and can lead to the development of metastasis (Vera et al. 2013; Müller and Bosserhoff 2008).

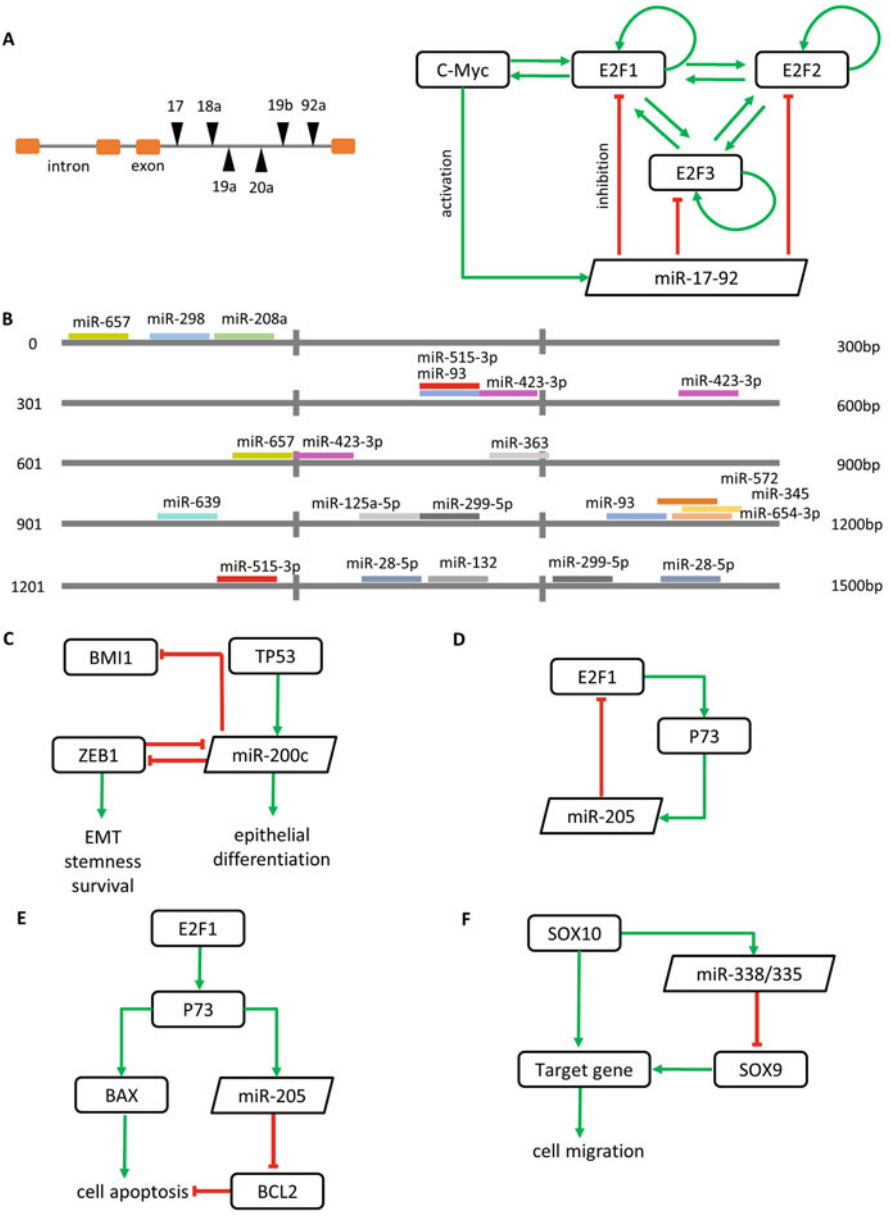


Fig. 1.5 A zoo of miRNA-mediated network motifs. (**a**) Example of a miRNA cluster. The miR-17-92 cluster composed of 6 miRNAs (indicated by arrows) is located in an intron of the miR-17-92 cluster host gene (MIR17HG). miR-17-92 cluster expression is regulated by the TF c-Myc. miR-17-92 cluster members repress transcripts of the E2F1 gene family. (**b**) An example of a miRNA target hub gene. The gene CDKN1A is a miRNA target hub with at least 22 predicted miRNA binding sites in its 3′ UTR. (**c**) An example of a positive feedback loop. ZEB1 and miR-200c repress the expression of one another, thereby forming a positive feedback loop. In addition, p53 activates transcription of miR-200c and miR-200c inhibits translation of stem cell

Feedforward loops Another frequently observed network motif involving miRNAs, TFs, and their mutual targets are feedforward loops (Lai and Vera 2013c). In such motifs, the TF regulates both the target gene and the miRNA, while the miRNA inhibits the mutual target. In this way, the TF regulates the target via two or more branches, i.e., directly via transcriptional regulation and indirectly via miRNA-mediated gene repression. While the notion of feedback loops is firmly imprinted in the experimentalist's way of thinking, feedforward loops have received little attention until recently (Jiang et al. 2019). We can distinguish two types of these loops: (1) a coherent feedforward loop, when the TF regulation is consistent through the two branches, (2) an incoherent feedforward loop when the TF regulation is inconsistent. *Coherent feedforward loops* can act as a safeguard mechanism, i.e., the effect on a downstream target is triggered only if both branches of the loop are active at the same time. This is the case for the feedforward loop established by E2F1, p73, and its apoptosis-related targets, which is mediated by miR-205 (Vera et al. 2013). Triggering apoptosis requires the expression of the pro-apoptotic targets of E2F1 and the coordinated repression of the anti-apoptotic protein BCL2 by miR-205 (Fig. 1.5e). Since both processes must occur simultaneously, this provides a window of opportunity for the cell to either confirm or prevent the irreversible activation of apoptosis. In the case of the E2F1-p73/DNp73-miR-205 circuit, miR-205 expression is inhibited via the oncogenic DNp73 splice isoform of p73 preventing apoptosis of some aggressive tumors. *Incoherent feedforward loops* can also induce sophisticated regulatory patterns. In recent publications (Reiprich et al. 2017; Cantone et al. 2019), researchers detected a plethora of feedforward loops linked to cell differentiation. Many of these motifs involve interactions between miRNAs and the SOX family, whose deregulation is critical in melanoma pathogenesis (Graf et al. 2014). One of the loops detected by Reiprich and his coworkers involves the TFs SOX10 and SOX9, the SOX9-repressing miR-338 and miR-335, and mutual targets of the TFs that promote cell migration. The authors hypothesized that the incoherent feedforward loop established (SOX10 → target → cell migration; SOX10 → miR-338/335 ⊣ SOX9 → target → cell migration) can generate a time window during differentiation, in which cell migration is possible because of the concerted activation of SOX10 and SOX9 (Fig. 1.5f). Beyond this period, SOX10-mediated activation of the two miRNAs represses SOX9 and the cells lose their migratory capacity. Sustained migratory capacity can be an advantage for metastatic tumors like melanoma and therefore the repression of this type of loop could be

Fig. 1.5 (continued) factors, such as BMI1. p53 stimulates expression of miR-200c, thereby driving epithelial differentiation and counteracting epithelial–mesenchymal transition (EMT) and stemness. (**d**) An example of a negative feedback loop. The TF E2F1 promotes the transcription of P73, which can upregulate the expression of miR-205. In turn, miR-205 represses E2F1, thereby forming a negative feedback loop. (**e**) An example of a coherent feedforward loop. Two signaling pathways can lead to upregulation of E2F1-related cell apoptosis—one via pro-apoptotic gene BAX and the other via anti-apoptotic gene BCL2 targeted by miR-205. (**f**) An example of an incoherent feedforward loop. SXO10 can regulate cell migration in an inconsistent manner—one pathway promotes it and the other suppresses it via miR-338 and miR-335

advantageous. Interestingly, miR-335 is repressed in metastatic melanoma (Cheng and Shen 2020), while SOX10 is overexpressed.

1.4 Bioinformatics and Systems Approaches as the "Lifeline" to Navigate miRNA Networks

As a single miRNA can have many targets, the effective regulation of its target genes may depend on other interacting molecules like lncRNAs and RNA binding proteins. Often, multiple miRNAs target the same central cancer gene or genes belonging to the same cancer pathway. The expression of each miRNA is regulated by different TFs, and miRNAs are entangled with their TFs and targets in feedback and feedforward loops. As a final point, these regulatory events do not happen in isolation, but they form large, densely connected regulatory networks of miRNAs, TFs, lncRNAs, and gene targets. The only way to navigate this level of complexity, gain insights into oncogenesis, and design personalized therapies is to develop and apply a systematic approach. Specifically, high-throughput molecular data (e.g., transcriptomes and proteomes) are analyzed and integrated utilizing bioinformatics algorithms and computational models. Bioinformatics algorithms that focus on the genome scale can be used to identify novel miRNA genes (Stegmayer et al. 2019), miRNA targets (Saçar Demirci et al. 2019), and mutations within miRNA binding sites (Ryan et al. 2016) or to detect miRNA clusters (Chan et al. 2012) and miRNA regulatory hubs (Mukherjee et al. 2019). Network-based analysis of omics data can be utilized to detect deregulated miRNA-mediated feedback and feedforward loops and to obtain core regulatory subnetworks important for the regulation of tumor initiation, progression, and therapy resistance (Dreyer et al. 2018). Finally, when considering the spatiotemporal dynamics of these circuits, one can utilize computational modeling (Vera et al. 2013; Lai et al. 2016). Ultimately, all these tools can be employed to obtain predictive gene signatures for cancer progression or stratification (Hayes et al. 2014) or to detect therapeutic miRNA targets (Lai et al. 2019).

Acknowledgments Our work in cancer and immunotherapy modeling is supported by the German Ministry of Education and Research (BMBF) through the e:Med initiatives: MelAutim [01ZX1905A to XL and JV] and KI-VesD [031L0244A to JV]. We also acknowledge the funding of our research activities in melanoma from the Manfred Roth-Stiftung, the Staedtler-Stiftung, the Trunk-Stiftung, the Matthias-Lackas-Stiftung, and the Hiege Stiftung. Financial support was also provided by the National Health & Medical Research Council Investigator Grant (#1196405) and by the Cancer Council NSW (RG20-12).

References

Alla V, Kowtharapu BS, Engelmann D, Emmrich S, Schmitz U, Steder M, Pützer BM (2012) E2F1 confers anticancer drug resistance by targeting ABC transporter family members and Bcl-2 via the p73/DNp73-miR-205 circuitry. Cell Cycle 11:3067–3078

Alon U (2007) Network motifs: theory and experimental approaches. Nat Rev Genet 8:450–461
Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116:281–297
Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136:215–233
Baumann V, Winkler J (2014) miRNA-based therapies: strategies and delivery platforms for oligonucleotide and non-oligonucleotide agents. Future Med Chem 6:1967–1984
Berezikov E (2011) Evolution of microRNA diversity and regulation in animals. Nat Rev Genet 12: 846–860
Bogusławska J, Rodzik K, Popławski P, Kędzierska H, Rybicka B, Sokół E, Tański Z, Piekiełko-Witkowska A (2018) TGF-β1 targets a microRNA network that regulates cellular adhesion and migration in renal cancer. Cancer Lett 412:155–169
Briskin D, Wang PY, Bartel DP (2020) The biochemical basis for the cooperative action of microRNAs. Proc Natl Acad Sci U S A 117:17764–17774
Burk U, Schubert J, Wellner U, Schmalhofer O, Vincan E, Spaderna S, Brabletz T (2008) A reciprocal repression between ZEB1 and members of the miR-200 family promotes EMT and invasion in cancer cells. EMBO Rep 9:582–589
Buschmann MD, Carrasco MJ, Alishetty S, Paige M, Alameh MG, Weissman D (2021) Nanomaterial delivery systems for mRNA vaccines. Vaccines (Basel) 9:65
Cantone M, Küspert M, Reiprich S, Lai X, Eberhardt M, Göttle P, Beyer F, Azim K, Küry P, Wegner M et al (2019) A gene regulatory architecture that controls region-independent dynamics of oligodendrocyte differentiation. Glia 67:825–843
Cardin S-E, Borchert GM (2017) Viral microRNAs, host microRNAs regulating viruses, and bacterial microRNA-like RNAs. Methods Mol Biol 1617:39–56
Chan W-C, Ho M-R, Li S-C, Tsai K-W, Lai C-H, Hsu C-N, Lin W (2012) MetaMirClust: discovery of miRNA cluster patterns using a data-mining approach. Genomics 100:141–148
Cheng Y, Shen P (2020) miR-335 acts as a tumor suppressor and enhances ionizing radiation-induced tumor regression by targeting ROCK1. Front Oncol 10:278
Czubak K, Lewandowska MA, Klonowska K, Roszkowski K, Kowalewski J, Figlerowicz M, Kozlowski P (2015) High copy number variation of cancer-related microRNA genes and frequent amplification of DICER1 and DROSHA in lung cancer. Oncotarget 6:23399–23416
Doench JG, Sharp PA (2004) Specificity of microRNA target selection in translational repression. Genes Dev 18:504–511
Dreyer FS, Cantone M, Eberhardt M, Jaitly T, Walter L, Wittmann J, Gupta SK, Khan FM, Wolkenhauer O, Pützer BM et al (2018) A web platform for the network analysis of high-throughput data in melanoma and its use to investigate mechanisms of resistance to anti-PD1 immunotherapy. Biochim Biophys Acta Mol basis Dis 1864:2315–2328
Dublanche Y, Michalodimitrakis K, Kümmerer N, Foglierini M, Serrano L (2006) Noise in transcription negative feedback loops: simulation and experimental analysis. Mol Syst Biol 2:41
Duran-Sanchon S, Moreno L, Augé JM, Serra-Burriel M, Cuatrecasas M, Moreira L, Martín A, Serradesanferm A, Pozo À, Costa R et al (2020) Identification and validation of MicroRNA profiles in fecal samples for detection of colorectal cancer. Gastroenterology 158:947–957.e4
Fabian MR, Sonenberg N, Filipowicz W (2010) Regulation of mRNA translation and stability by microRNAs. Annu Rev Biochem 79:351–379
Filipowicz W, Bhattacharyya SN, Sonenberg N (2008) Mechanisms of post-transcriptional regulation by microRNAs: are the answers in sight? Nat Rev Genet 9:102–114
Forloni M, Dogra SK, Dong Y, Conte D, Ou J, Zhu LJ, Deng A, Mahalingam M, Green MR, Wajapeyee N (2014) miR-146a promotes the initiation and progression of melanoma by activating Notch signaling. elife 3:e01460
Foulkes WD, Priest JR, Duchaine TF (2014) DICER1: mutations, microRNAs and mechanisms. Nat Rev Cancer 14:662–672
Friedman RC, Farh KK-H, Burge CB, Bartel DP (2009) Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 19:92–105
Fuziwara CS, Kimura ET (2015) Insights into regulation of the miR-17-92 cluster of miRNAs in cancer. Front Med (Lausanne) 2:64

Gangaraju VK, Lin H (2009) MicroRNAs: key regulators of stem cells. Nat Rev Mol Cell Biol 10: 116–125

Ghousein A, Mosca N, Cartier F, Charpentier J, Dupuy J-W, Raymond A-A, Bioulac-Sage P, Grosset CF (2020) miR-4510 blocks hepatocellular carcinoma development through RAF1 targeting and RAS/RAF/MEK/ERK signalling inactivation. Liver Int 40:240–251

Glinge C, Clauss S, Boddum K, Jabbari R, Jabbari J, Risgaard B, Tomsits P, Hildebrand B, Kääb S, Wakili R et al (2017) Stability of circulating blood-based microRNAs - pre-analytic methodological considerations. PLoS One 12:e0167969

Graf SA, Busch C, Bosserhoff A-K, Besch R, Berking C (2014) SOX10 promotes melanoma cell invasion by regulating melanoma inhibitory activity. J Invest Dermatol 134:2212–2220

Guo L, Bai Y, Ji S, Ma H (2019) MicroRNA-98 suppresses cell growth and invasion of retinoblastoma via targeting the IGF1R/k-Ras/Raf/MEK/ERK signaling pathway. Int J Oncol 54:807–820

Hata A, Kashima R (2016) Dysregulation of microRNA biogenesis machinery in cancer. Crit Rev Biochem Mol Biol 51:121–134

Hayes J, Peruzzi PP, Lawler S (2014) MicroRNAs in cancer: biomarkers, functions and therapy. Trends Mol Med 20:460–469

Hirano D, Hayakawa F, Yasuda T, Tange N, Yamamoto H, Kojima Y, Morishita T, Imoto N, Tsuzuki S, Mano H et al (2019) Chromosomal translocation-mediated evasion from miRNA induces strong MEF2D fusion protein expression, causing inhibition of PAX5 transcriptional activity. Oncogene 38:2263–2274

Hoffmann A, Natoli G, Ghosh G (2006) Transcriptional regulation via the NF-kappaB signaling module. Oncogene 25:6706–6716

Hwang H-W, Mendell JT (2006) MicroRNAs in cell proliferation, cell death, and tumorigenesis. Br J Cancer 94:776–780

Inoue K, Shinohara H, Behar M, Yumoto N, Tanaka G, Hoffmann A, Aihara K, Okada-Hatakeyama M (2016) Oscillation dynamics underlie functional switching of NF-κB for B-cell activation. NPJ Syst Biol Appl 2:16024

Ishida M, Selaru FM (2013) miRNA-based therapeutic strategies. Curr Anesthesiol Rep 1:63–70

Jia D, Li L, Andrew S, Allan D, Li X, Lee J, Ji G, Yao Z, Gadde S, Figeys D et al (2017) An autocrine inflammatory forward-feedback loop after chemotherapy withdrawal facilitates the repopulation of drug-resistant breast cancer cells. Cell Death Dis 8:e2932

Jiang L, Yu X, Ma X, Liu H, Zhou S, Zhou X, Meng Q, Wang L, Jiang W (2019) Identification of transcription factor-miRNA-lncRNA feed-forward loops in breast cancer subtypes. Comput Biol Chem 78:1–7

Jinesh GG, Flores ER, Brohl AS (2018) Chromosome 19 miRNA cluster and CEBPB expression specifically mark and potentially drive triple negative breast cancers. PLoS One 13:e0206008

Kozomara A, Birgaoanu M, Griffiths-Jones S (2019) miRBase: from microRNA sequences to function. Nucleic Acids Res 47:D155–D162

Krol J, Loedige I, Filipowicz W (2010) The widespread regulation of microRNA biogenesis, function and decay. Nat Rev Genet 11:597–610

Kumar MS, Pester RE, Chen CY, Lane K, Chin C, Lu J, Kirsch DG, Golub TR, Jacks T (2009) Dicer1 functions as a haploinsufficient tumor suppressor. Genes Dev 23:2700–2704

Lai X, Vera J (2013a) MicroRNA clusters. In: Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H (eds) Encyclopedia of systems biology. Springer, New York, pp 1310–1314

Lai X, Vera J (2013b) Target hub. In: Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H (eds) Encyclopedia of systems biology. Springer, New York, pp 2134–2138

Lai X, Vera J (2013c) MicroRNA regulation, feed-forward loops. In: Dubitzky W, Wolkenhauer O, Cho K-H, Yokota H (eds) Encyclopedia of systems biology. Springer, New York, pp 1324–1328

Lai X, Schmitz U, Gupta SK, Bhattacharya A, Kunz M, Wolkenhauer O, Vera J (2012a) Computational analysis of target hub gene repression regulated by multiple and cooperative miRNAs. Nucleic Acids Res 40:8818–8834

Lai X, Wolkenhauer O, Vera J (2012b) Modeling miRNA regulation in cancer signaling systems: miR-34a regulation of the p53/Sirt1 signaling module. Methods Mol Biol 880:87–108
Lai X, Bhattacharya A, Schmitz U, Kunz M, Vera J, Wolkenhauer O (2013) A systems' biology approach to study microRNA-mediated gene regulatory networks. Biomed Res Int 2013:703849
Lai X, Wolkenhauer O, Vera J (2016) Understanding microRNA-mediated gene regulatory networks through mathematical modelling. Nucleic Acids Res 44:6019–6035
Lai X, Gupta SK, Schmitz U, Marquardt S, Knoll S, Spitschak A, Wolkenhauer O, Pützer BM, Vera J (2018) MiR-205-5p and miR-342-3p cooperate in the repression of the E2F1 transcription factor in the context of anticancer chemotherapy resistance. Theranostics 8:1106–1120
Lai X, Eberhardt M, Schmitz U, Vera J (2019) Systems biology-based investigation of cooperating microRNAs as monotherapy or adjuvant therapy in cancer. Nucleic Acids Res 47:7753–7766
Lake D, Corrêa SAL, Müller J (2016) Negative feedback regulation of the ERK1/2 MAPK pathway. Cell Mol Life Sci 73:4397–4413
Le MTN, Teh C, Shyh-Chang N, Xie H, Zhou B, Korzh V, Lodish HF, Lim B (2009) MicroRNA-125b is a novel negative regulator of p53. Genes Dev 23:862–876
Lin S, Gregory RI (2015) MicroRNA biogenesis pathways in cancer. Nat Rev Cancer 15:321–333
Lytle JR, Yario TA, Steitz JA (2007) Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5' UTR as in the 3′ UTR. Proc Natl Acad Sci U S A 104:9667–9672
Mao Z, Zhao H, Qin Y, Wei J, Sun J, Zhang W, Kang Y (2020) Post-transcriptional dysregulation of microRNA and alternative polyadenylation in colorectal cancer. Front Genet 11:64
Mastroianni J, Stickel N, Andrlova H, Hanke K, Melchinger W, Duquesne S, Schmidt D, Falk M, Andrieux G, Pfeifer D et al (2019) miR-146a controls immune response in the melanoma microenvironment. Cancer Res 79:183–195
Miyoshi K, Miyoshi T, Siomi H (2010) Many ways to generate microRNA-like small RNAs: non-canonical pathways for microRNA production. Mol Gen Genomics 284:95–103
Moszyńska A, Gebert M, Collawn JF, Bartoszewski R (2017) SNPs in microRNA target sites and their potential role in human disease. Open Biol 7:170019
Mukherjee S, Akbar I, Bhagat R, Hazra B, Bhattacharyya A, Seth P, Roy D, Basu A (2019) Identification and classification of hubs in microRNA target gene networks in human neural stem/progenitor cells following Japanese encephalitis virus infection. mSphere 4:e00588–e00519
Müller DW, Bosserhoff A-K (2008) Integrin beta 3 expression is regulated by let-7a miRNA in malignant melanoma. Oncogene 27:6698–6706
Ørom UA, Nielsen FC, Lund AH (2008) MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. Mol Cell 30:460–471
Piotto C, Julier Z, Martino MM (2018) Immune regulation of tissue repair and regeneration via miRNAs-new therapeutic target. Front Bioeng Biotechnol 6:98
Poma AM, Condello V, Denaro M, Torregrossa L, Elisei R, Vitti P, Basolo F (2019) DICER1 somatic mutations strongly impair miRNA processing even in benign thyroid lesions. Oncotarget 10:1785–1797
Quackenbush JF, Cassidy PB, Pfeffer LM, Boucher KM, Hawkes JE, Pfeffer SR, Kopelovich L, Leachman SA (2014) Isolation of circulating microRNAs from microvesicles found in human plasma. Methods Mol Biol 1102:641–653
Raimo M, Orso F, Grassi E, Cimino D, Penna E, De Pittà C, Stadler MB, Primo L, Calautti E, Quaglino P et al (2016) miR-146a exerts differential effects on melanoma growth and metastatization. Mol Cancer Res 14:548–562
Reiprich S, Cantone M, Weider M, Baroti T, Wittstatt J, Schmitt C, Küspert M, Vera J, Wegner M (2017) Transcription factor Sox10 regulates oligodendroglial Sox9 levels via microRNAs. Glia 65:1089–1102
Ruby JG, Jan CH, Bartel DP (2007) Intronic microRNA precursors that bypass Drosha processing. Nature 448:83–86
Ryan BC, Werner TS, Howard PL, Chow RL (2016) ImiRP: a computational approach to microRNA target site mutation. BMC Bioinformatics 17:190

Saçar Demirci MD, Yousef M, Allmer J (2019) Computational prediction of functional MicroRNA-mRNA interactions. Methods Mol Biol 1912:175–196
Saetrom P, Heale BSE, Snøve O, Aagaard L, Alluin J, Rossi JJ (2007) Distance constraints between microRNA target sites dictate efficacy and cooperativity. Nucleic Acids Res 35:2333–2342
Sapre N, Macintyre G, Clarkson M, Naeem H, Cmero M, Kowalczyk A, Anderson PD, Costello AJ, Corcoran NM, Hovens CM (2016) A urinary microRNA signature can predict the presence of bladder urothelial carcinoma in patients undergoing surveillance. Br J Cancer 114:454–462
Schmitz U, Lai X, Winter F, Wolkenhauer O, Vera J, Gupta SK (2014) Cooperative gene regulation by microRNA pairs and their identification using a computational workflow. Nucleic Acids Res 42:7539–7552
Shalgi R, Lieber D, Oren M, Pilpel Y (2007) Global and local architecture of the mammalian microRNA-transcription factor regulatory network. PLoS Comput Biol 3:e131
Soh J, Cho H, Choi C-H, Lee H (2018) Identification and characterization of MicroRNAs associated with somatic copy number alterations in cancer. Cancers (Basel) 10:E475
Stegeman S, Moya L, Selth LA, Spurdle AB, Clements JA, Batra J (2015) A genetic variant of MDM4 influences regulation by multiple microRNAs in prostate cancer. Endocr Relat Cancer 22:265–276
Stegmayer G, Di Persia LE, Rubiolo M, Gerard M, Pividori M, Yones C, Bugnon LA, Rodriguez T, Raad J, Milone DH (2019) Predicting novel microRNA: a comprehensive comparison of machine learning approaches. Brief Bioinform 20:1607–1620
Van Laar R, Lincoln M, Van Laar B (2018) Development and validation of a plasma-based melanoma biomarker suitable for clinical use. Br J Cancer 118:857–866
Vandenwijngaert S, Ledsky CD, Agha O, Wu C, Hu D, Bagchi A, Domian IJ, Buys ES, Newton-Cheh C, Bloch DB (2018) MicroRNA-425 and microRNA-155 cooperatively regulate atrial natriuretic peptide expression and cGMP production. PLoS One 13:e0196697
Vella MC, Choi E-Y, Lin S-Y, Reinert K, Slack FJ (2004) The C. elegans microRNA let-7 binds to imperfect let-7 complementary sites from the lin-41 3′UTR. Genes Dev 18:132–137
Vera J, Schmitz U, Lai X, Engelmann D, Khan FM, Wolkenhauer O, Pützer BM (2013) Kinetic modeling-based detection of genetic signatures that provide chemoresistance via the E2F1-p73/DNp73-miR-205 network. Cancer Res 73:3511–3524
Walz AL, Ooms A, Gadd S, Gerhard DS, Smith MA, Guidry Auvil JM, Guidry Auvil JM, Meerzaman D, Chen Q-R, Hsu CH et al (2015) Recurrent DGCR8, DROSHA, and SIX homeodomain mutations in favorable histology Wilms tumors. Cancer Cell 27:286–297
Wilking MJ, Singh C, Nihal M, Zhong W, Ahmad N (2014) SIRT1 deacetylase is overexpressed in human melanoma and its small molecule inhibition imparts anti-proliferative response via p53 activation. Arch Biochem Biophys 563:94–100
Winkle M, El-Daly SM, Fabbri M, Calin GA (2021) Noncoding RNA therapeutics - challenges and potential solutions. Nat Rev Drug Discov. https://doi.org/10.1038/s41573-021-00219-z
Wu S, Huang S, Ding J, Zhao Y, Liang L, Liu T, Zhan R, He X (2010) Multiple microRNAs modulate p21Cip1/Waf1 expression by directly targeting its 3′ untranslated region. Oncogene 29:2302–2308
Xiao M, Li J, Li W, Wang Y, Wu F, Xi Y, Zhang L, Ding C, Luo H, Li Y et al (2017) MicroRNAs activate gene transcription epigenetically as an enhancer trigger. RNA Biol 14:1326–1334
Yamakuchi M, Ferlito M, Lowenstein CJ (2008) miR-34a repression of SIRT1 regulates apoptosis. Proc Natl Acad Sci U S A 105:13421–13426
Yi C, Wang Q, Wang L, Huang Y, Li L, Liu L, Zhou X, Xie G, Kang T, Wang H et al (2012) MiR-663, a microRNA targeting p21 WAF1/CIP1, promotes the proliferation and tumorigenesis of nasopharyngeal carcinoma. Oncogene 31:4421–4433
Zhang H, Chen Y, Chen Y (2012) Noise propagation in gene regulation networks involving interlinked positive and negative feedback loops. PLoS One 7:e51840
Zhang L, Liao Y, Tang L (2019) MicroRNA-34 family: a potential tumor suppressor and therapeutic candidate in cancer. J Exp Clin Cancer Res 38:53

Chapter 2
Circulating MicroRNAs as Cancer Biomarkers in Liquid Biopsies

Beatriz Suárez, Carla Solé, Maitane Márquez, Francesca Nanetti, and Charles Henderson Lawrie

Abstract Biological fluids such as blood, saliva, and urine offer a rich source of biomarkers that have the potential to change the paradigm of cancer management. By allowing routine noninvasive sampling that can offer new insights into cancer progression and response to treatment so-called liquid biopsies can play an important role in personalized medicine. MicroRNAs (miRNAs) are a particularly attractive class of biomarkers as they are not only resistant to the high levels of RNases found in biological fluids, but also able to confer clinically useful information about the disease relating to diagnosis, prognosis, and the response to treatment. Circulating miRNAs are either associated with proteins or extracellular vesicles (EV) and although their origin and implied functions as intracellular messengers remain somewhat controversial, they are implicated in the progression and the establishment of metastatic niches. In this chapter, we review the rapidly emerging field of circulating miRNA cancer biomarkers, their origin and function, techniques to detect these molecules, and the use of bioinformatic tools to derive implied regulatory function, as well as the challenges that lie ahead for their clinical implementation.

Keywords Biomarkers · Cancer · ncRNAs · miRNAs · Liquid biopsy · Bioinformatics

Beatriz Suárez and Carla Solé contributed equally with all other contributors.

B. Suárez · C. Solé · M. Márquez · F. Nanetti
Molecular Oncology Group, Biodonostia Research Institute, San Sebastián, Spain

C. H. Lawrie (✉)
Molecular Oncology Group, Biodonostia Research Institute, San Sebastián, Spain

IKERBASQUE, Basque Foundation for Science, Bilbao, Spain

Radcliffe Department of Medicine, University of Oxford, Oxford, UK

Sino-Swiss Institute of Advanced Technology, Shanghai University, Shanghai, China
e-mail: charles.lawrie@biodonostia.org

U. Schmitz et al. (eds.), *Systems Biology of MicroRNAs in Cancer*, Advances in Experimental Medicine and Biology 1385,
https://doi.org/10.1007/978-3-031-08356-3_2

Abbreviations

ABs	Apoptotic bodies
AC	Adjuvant chemotherapy
ADC	Adenocarcinoma
AFP	Alpha-Fetoprotein
Ago1/2	Argonaute 1 or 2
AGO CLIP-seq	Argonaute-crosslinking and immunoprecipitation sequencing
ALIX	ALG-2-interacting protein X
AMPS	Acute multiple sclerosis plaques
ARF6	(Adenosine diphosphate)-ribosylation factor 6
ATG7	Autophagy-related 7 protein
ATG12	Autophagy-related 12 protein
ATP6V1E1	V-Type proton ATPase subunit E 1
AUC	Area under the curve
BC	Breast cancer
BCa	Bladder cancer
BCR	Biochemical recurrence
BCT	Benign colorectal tumor
BFS	Breadth first search
BGD	Benign gastric disease
BI-RADS	Breast imaging reporting and data system
BM	Bone marrow
Biomark.	Biomarker
BPH	Benign prostate hyperplasia
C	Chemotherapy
CA19–9	Carbohydrate antigen 19–9
CA72–4	Cancer antigen 72–4
CAG	Chronic atrophic gastritis
CAGE	Cap analysis of gene expression
Cancer GAMAdb	Cancer genome-wide association and meta-analyses database
CC	Cervical cancer
ccRCC	Clear cell renal cell carcinoma
CDB	Colorectal benign disease
CDS	Coding sequence
CEA	Carcinoembryonic antigen
cfRNA	Circulating free RNA
CHB	Chronic hepatitis B
ChIP	Chromatin immunoprecipitation
CIN	Cervical intraepithelial neoplasia
circRNA	circular RNA
CLASH	Cross-linking, ligation and sequencing of hybrids
CLEAR-CLIP	Covalent-ligation of endogenous ago-bound RNAs-CLIP sequencing

CLIP-seq	Cross-linking immunoprecipitation sequencing
CLL	Chronic lymphocytic leukemia
CMEP	Circulating microRNA expression profiling
CNS	Central nervous system
COPD	Chronic obstructive pulmonary disease
COSMIC	Catalogue of somatic mutations in cancer
CR	Chemoresistant
CRC	Colorectal cancer
CRPs	Ceramide-rich platforms
CS	Chemosensitive
CSD	Cold-shock domain
CSF	Cerebro spinal fluid
CT	Computed tomography
CTCs	Circulating tumor cells
CtDNA	Circulating tumor-associated DNA
CtRNA	Circulating tumor-associated RNA
D	Diagnostic
DAVID	Database for annotation, visualization and integrated discovery
dbDEMC	Database of differentially expressed miRNAs in human cancers
DbGaP	Database of Genotypes and Phenotypes
DDBJ	DNA Data Bank of Japan
ddPCR	Droplet digital PCR
DFS	Disease-free survival
DgC	Digestive cancer
DLBCL	Diffuse large B-cell lymphoma
EC	Endometrial cancer
EFS	Event-free survival
ELISA	Enzyme-linked immunosorbent assay
ENCODE	Encyclopedia of DNA elements
EOC	Epithelial ovarian cancer
ESCC	Esophageal squamous cell carcinoma
ESCRT	Endosomal sorting complexes required for transport
EV	Extracellular vesicle
Exp	Expression
Exo	Exosome
FANTOM	Functional annotation of the mammalian genome
FBLs	Feed-back loops
FDA-NIH	Food and Drug Administration and National Institute of Health
FFLs	Feed-forward loops
FMR1	Fragile X Mental Retardation
FOLFIRI	Folinic acid, fluorouracil and irinotecan
FOLFOX	5-Fluorouracil, leucovorin, and oxaliplatin
FPC	Familial pancreatic cancer
GBM	Glioblastoma multiforme

G3BP1	GTPase activating protein (SH3 Domain) binding protein 1
GC	Gastric cancer
GEO	Gene expression omnibus
GO	Gene ontology
GPL	Gastric precancerous lesions
GRO/PRO-Seq	Global and precision nuclear run-on sequencing
GW182	Glycine-tryptophan protein of 182 kDa
GWAS	Genome-Wide Association Study
H3K4me3	Trimethylation at the fourth lysine residue of the histone H3 protein
H3K36me3	Tri-methylation at the 36th lysine residue of the histone H3 protein
H3K27Ac	Acetylation at the 27th lysine residue of the histone H3 protein
H3K4me1	Monomethylation at the fourth lysine residue of the histone H3 protein
HC	Healthy control
HCC	Hepatocellular carcinoma
HDL	High-density lipoprotein
hEXO	GGCU/A sequence
HITS-CLIP	High-throughput sequencing of RNA isolated by crosslinking immunoprecipitation
HMDD	Human MicroRNA Disease Database
HnRNPA2B1	Heterogeneous nuclear ribonucleoproteins A2/B1
HnRNPK	Heterogeneous nuclear ribonucleoprotein K
HnRNPU	Heterogeneous nuclear ribonucleoprotein U
HNSCC	Head and neck squamous cell carcinoma
HOTTIP	HOXA distal transcript antisense RNA
Hp	Helicobacter pylori
Hrs	Hepatocyte growth factor [HGF]-regulated tyrosine kinase substrate
HOTAIR	Hox transcript antisense intergenic RNA
HSP70/90	Heat shock protein 70 or 90
IGM	Idiopathic granulomatous mastitis
ILVs	Intraluminal vesicles
INDELs	Insertion/Deletion
IP-seq	Immunoprecipitation sequencing
ITGαvβ	Integrin subunit beta 1
KEGG	Kyoto Encyclopedia of Genes and Genomes
KPS	Karnofsky performance scale
KRAS	Kirsten Rat Sarcoma Viral (Proto)-Oncogene
LARP1	La ribonucleoprotein 1
LC	Lung cancer
LC3	Microtubule-associated protein 1 light chain 3 alpha
LDELS	LC3-dependent EV loading and secretion

LDL	Low-density lipoproteins
LIR	LC3 interaction region
lncRNA	Long non-coding RNA
LUAD	Lung adenocarcinoma
MALAT1	Metastasis associated lung adenocarcinoma transcript 1
MDS	Myelodysplastic syndrome
MEG3	Maternally expressed 3
MGUS	Monoclonal gammopathy of undetermined significance
MIAME	Minimum Information About a Microarray Experiment
MiRNA	MicroRNA
MRI	Magnetic resonance imaging
MS	Median survival
MVBs	Multivesicular bodies
ncRNAs	non-coding RNAs
NEAT1	Nuclear enriched abundant transcript 1
NHGRI	The National Human Genome Research Institute Catalog
NGS	Next generation sequencing
NN	Neural network
NRAS	Neuroblastoma RAS viral oncogene homolog
NSCLC	Non-small cell lung cancer
nSMAse2	Neutral sphingomyelinase 2
NURR	N-terminal unit for RNA recognition
OC	Ovarian cancer
OS	Overall survival
OS	Osteosarcoma
OSCC	Oral squamous cell carcinoma
PA	Pituitary adenoma
PAR-CLIP	Photoactivatable ribonucleoside-enhanced crosslinking and immunoprecipitation
PC	Pancreatic cancer
PCa	Prostate cancer
PCAT1	Prostate cancer associated transcript 1
PCNSL	Primary central nervous system lymphoma
Pd	Predictive
PD	Progressive disease
PDAC	Pancreatic ductal adenocarcinoma
PD-1	Programmed death 1
PD-L1	Programmed death-ligand 1
PE	Phosphatidylethanolamide
PET	Positron emission tomography
PFS	Progression-free survival
Pg	Prognostic
PID	Pathway Interaction Database
piRNA	Piwi-associated RNAs

POM1	1-Month post-operation
POM6	6-Month post-operation
PR	Partial response
PS	Phosphatidylserine
PSA	Prostate-specific antigen
PTENP1	Phosphatase and tensin homolog pseudogene 1
PVT1	Plasmacytoma variant translocation 1
R	Responders
RBP	RNA-binding proteins
RCC	Renal cell carcinoma
Ref	Reference
RIP-Seq	RNA immunoprecipitation sequencing
RISC	RNA-induced silencing complex
RFS	Recurrence-free survival
RILP	Rab interacting lysosomal protein
RPF-Seq	Ribosome profiling sequencing
RRM	RNA recognition motif
RT-qPCR	Quantitative reverse transcription PCR
SACC	Salivary adenoid cystic carcinoma
SAFB	Scaffold attachment factor B
SCC	Squamous cell carcinoma
SCCA	Squamous cell carcinoma antigen
SCLC	Small cell lung cancer
SF3A1	Splicing factor 3a subunit 1
Sm	Smoker
SNP	Single nucleotide polymorphism
SnRNA	Small nuclear RNA
SnoRNA	Small nucleolar RNA
SPRY4-IT1	SPRY4 intronic transcript 1
SRA	Short read archive
SR-BI	Scavenger receptor class B type I
SVM	Support vector machine
SYNCRIP	Synaptotagmin-binding cytoplasmic RNA-interacting protein
TCGA	The Cancer Genome Atlas
TDEs	Tumor-derived exosomes
TFs	Transcription factors
TKI	Tyrosine kinase inhibitor
TLR8	Toll-like receptor 8
TME	Tumor microenvironment
TMVs	Tumor-derived microvesicles
TNM	Tumor-node-metastasis
TRBP	Transactivation response RNA-binding protein
TRNAs	transfer RNA
TSG101	Tumor susceptibility gene 101 protein

TSSs	Transcriptional start sites
UCA1	Urothelial cancer associated 1
UCB	Urotherial carcinoma of the bladder
UTR	Untranslated region
US	Urinary sediment
VCAMP3	Vascular cell adhesion molecule
XELOX	Capecitabine plus oxaliplatin
YBX1	Y box binding protein 1

2.1 Introduction

Cancer is the primary cause of death before the age of 70 years in 91 out of 172 countries worldwide, and there are more than 18 million new cases and 9 million mortalities caused by cancer every single year (Bray et al. 2018). Early diagnosis of cancer is a critical factor in improving the survival of cancer patients, that in turn depends upon the availability of suitably robust biomarkers and the widespread technology to detect these biomarkers. Biomarkers are defined by the FDA-NIH as "a defined characteristic that is measured as an indicator of normal biological processes, pathogenic processes, or responses to an exposure or intervention, including therapeutic interventions" (Group F-NBW 2016; Califf 2018). In general, cancer biomarkers can be categorized as (1) diagnostic biomarkers, to detect whether a patient has a particular medical condition; (2) prognostic biomarkers that indicate an increased (or decreased) likelihood of a future clinical event, disease recurrence, or progression in an identified population; and (3) predictive biomarkers that measure the likelihood of response or lack of response for a particular therapy. There are many technologies available for the detection of biomarkers at both research and clinical levels. In the cancer clinic, this can involve medical history file review, physical examination, diagnostic procedures (e.g., bronchoscopy, endoscopy, and lumbar punch), imaging tests (e.g., X-ray, PET, CT, MRI, and ultrasound), histological examination, biochemical analysis of blood or urine, and increasingly the use of genetic and molecular tests, most commonly carried out on biopsy material but increasingly in biological fluids (i.e., liquid biopsies).

Biological fluids, most commonly blood, but also plasma, urine, pleural fluid, cerebrospinal fluid (CSF), and saliva, represents a rich source of nucleic acid biomarkers (Fig. 2.1). These can come from circulating tumor cells (CTCs), extracellular vesicles (EV) or be free-floating, and include circulating tumor-associated DNA (ctDNA), circulating tumor-associated RNA (ctRNA), microRNAs (miRNAs), and other noncoding RNAs (ncRNAs) (Alix-Panabières and Pantel 2016; Schwarzenbach et al. 2014; Panagiotara et al. 2017; Michela 2021; Sole et al. 2019).

MiRNAs are single-strand small (20–23 nucleotides length) noncoding RNAs that posttranscriptionally regulate gene expression (Rupaimoole and Slack 2017). They constitute 1% of the human genome and regulate more than 50% of protein-coding genes in mammals (Hsu et al. 2006). MiRNAs bind to complementary or near

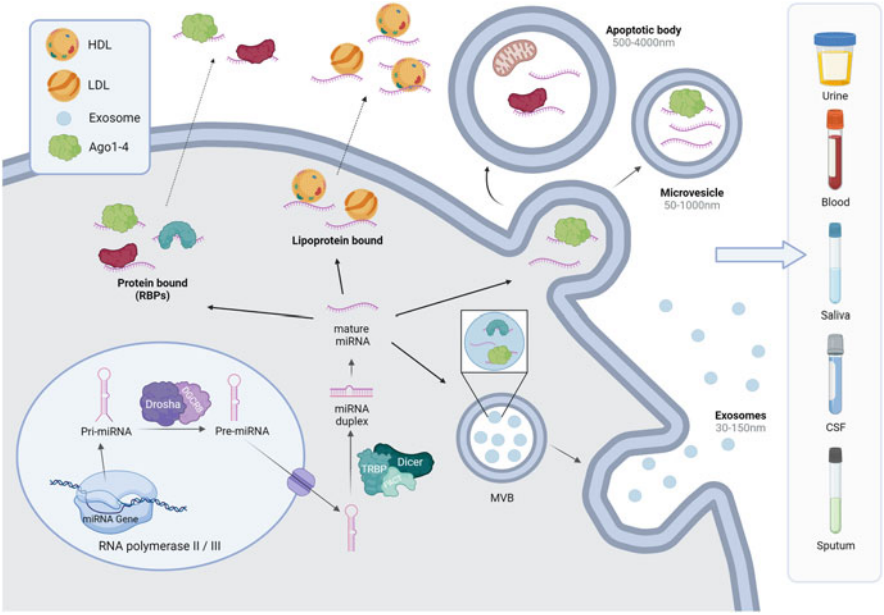


Fig. 2.1 MicroRNA synthesis and release into circulation. MiRNA genes are transcribed by RNA Polymerase II and III into pri-miRNAs, which are processed by Drosha and DGCR8 enzymes to produce pre-miRNAs. The pre-miRNAs are exported to the cytoplasm, where they are cleaved by the action of Dicer/TRBP/PACT complex to form the miRNA duplex intermediate. Then, the interaction with Ago and the RISC complex allows the final maturation of miRNAs. Apart from other passive release mechanisms, miRNAs can then be delivered out of the donor cells thanks to the binding to RNA-binding proteins (RBPs), lipoproteins, or encapsulation within extracellular vesicles. The extracellular vesicles can be generated either through the shedding of the plasma membrane or the formation of multivesicular bodies containing intraluminal vesicles that fuse with the plasma membrane. The released miRNAs can be detected in different human fluids including urine, blood, saliva, cerebrospinal fluid (CSF), or sputum and therefore be used as biomarkers

complementary sequences in the 3′-untranslated regions (3′-UTR) of their messenger RNA (mRNA) targets, and occasionally to 5′-UTR or coding sequences. They act primarily by post-transcriptional reduction of translation or through direct degradation of their target mRNA (Bartel 2004), although other noncanonical pathways also exist (Cullen 2004).

Although the gold standard for cancer diagnosis remains the histological examination of tissue biopsy, such procedures are expensive, sometimes risky, and require consistent evaluation by expert pathologists. Liquid biopsies are an attractive alternative as their noninvasive nature allows for rapid, economical, and repeated sampling, characteristics that would allow their use in screening programs and the monitorization of treatment response and disease progression for a more personalized approach to cancer management. Furthermore, the use of core tissue biopsies, in particular, has been called into question recently as intra-tumoral heterogeneity can be associated with false negatives, potentially leading to undertreatment of patients,

especially when biomarker expression is low, such as is the case for the widely used immunotherapy biomarker PD-L1 (López et al. 2018). Liquid biopsies, in contrast, can capture the entire genetic panorama of the tumoral landscape.

The history of circulating nucleic acids dates back to the 1940s, when Mandel and Mètais identified DNA and RNA in the plasma of healthy individuals (Mandel and Metais 1948). This remained something of an oddity until 1994 when circulating *KRAS* DNA mutations were identified in the plasma of pancreatic cancer (PC) patients and, in another study, circulating *NRAS* mutations were identified in myelodysplastic syndrome (MDS) patients, opening the possibility of their use as cancer biomarkers (Sorenson et al. 1994; Vasioukhin et al. 1994). The first mention of circulating free RNA (cfRNA) came 5 years later in 1999, when it was identified in the blood of nasopharyngeal carcinoma patients (Lo et al. 1999). In 2008, circulating miRNAs were identified in the sera of diffuse large B-cell lymphoma (DLBCL) patients, and a few months later in the sera of prostate cancer (PCa) patients (Lawrie et al. 2008; Mitchell et al. 2008).

Circulating miRNAs, in particular, have generated much interest in recent years as in addition to the advantages listed above for circulating biomarkers, the small size of circulating miRNAs not only makes them much more stable than other RNA types, but they are also present in extracellular vesicles providing a potentially tumor-enriched source of biomarkers. In this chapter, we consider the accumulating evidence that circulating miRNAs are useful cancer biomarkers and the challenges that remain before they enter routine clinical practice.

2.2 Circulating miRNAs

The labile nature of RNA and the presence of RNases at high concentrations in biological fluids has been well-known for many years (Duttagupta et al. 2011). It was therefore somewhat surprising when miRNAs were identified in blood, and furthermore that they were stable and protected from RNase degradation (Mitchell et al. 2008). Subsequent studies that used detergents, proteases, and sonication suggested that miRNAs are not inherently resistant to RNase degradation but are, in fact, protected by their association with a carrier protein or lipid, either associated with extracellular vesicles or free-floating (Mitchell et al. 2008; Chen et al. 2008; Kosaka et al. 2010a).

MiRNAs appear to be ubiquitously present in biological fluids and have been detected not only in blood but also in plasma, serum, saliva, urine, tears, semen, cerebral spinal fluid, milk, and ocular fluid (Weber et al. 2010; Fleischhacker and Schmidt 2010). The origin of circulating miRNAs remains somewhat controversial and there are at least three different non-mutually exclusive mechanisms that have been identified (Turchinovich et al. 2012; Chatrchyan et al. 2011). Firstly, the passive release of miRNAs from broken cells and tissues following injury, chronic inflammation, cell apoptosis or necrosis, or from cells with a short half-life, such as platelets. Secondly, by the active secretion of miRNAs contained within cellular

fragments of EVs such as exosomes, microvesicles, and microparticles (Mitchell et al. 2008; Aguda et al. 2008; Ismail et al. 2013; Valadi et al. 2007a). Thirdly, by active secretion of miRNAs associated with RNA-binding proteins (RBP) forming conjugated complexes such as lipoproteins (Vickers et al. 2011), nucleophosmin (Wang et al. 2010; Aad et al. 2010), or Argonaute 1 and 2 (Ago1 and Ago2) (Arroyo et al. 2011; Turchinovich et al. 2011) (Fig. 2.1). Among lipoproteins, the most commonly associated with circulating miRNAs are High-Density Lipoproteins (HDL) and, to a lesser extent Low-Density Lipoproteins (LDL) (Wagner et al. 2013). The Scavenger Receptor class B type I (SR-BI) appears necessary for the uptake of HDL–miRNA complexes on recipient cells and to prevent the degradation at the lysosomes once they have been internalized in the cell (Vickers et al. 2011).

In addition to being free-floating, RBPs can also be present within EVs, representing up to 25% of the protein content (Sork et al. 2018). There are several types of EVs that are classified according to size along with their respective mechanism of origin (Fig. 2.1). Exosomes, for example, are vesicles of 30–150 nm in size generated by invagination of the plasma membrane forming intraluminal vesicles (ILVs) within multivesicular bodies (MVBs) (Johnstone et al. 1987; Raposo et al. 1996; Colombo et al. 2014). Exosomes can be differentiated from other EVs due to their sedimentation by ultracentrifugation at 100,000 g, the expression of specific surface markers such as tetraspanins (CD63, CD81, and CD9), flotillin, as well as an enrichment in Rab GTPases, ALIX, TSG101, HSP70, and HSP90 proteins in the lumen (Escola et al. 1998; Buschow et al. 2009; Tauro et al. 2012; Wubbolts et al. 2003). On the other hand, microvesicles (MVs) are vesicles of 50–1000 nm in size that are generated by direct budding and fission of the plasma membrane (Cocucci et al. 2009). MVs can be differentiated from other EVs because they sediment at 10,000–60,000 g and also express specific surface markers including β-integrins, CD40, P-selectin (CD62), ARF6, Annexin A1, and VCAMP3 (Heijnen et al. 1999; Jeppesen et al. 2019; Muralidharan-Chari et al. 2009). Apoptotic bodies (ABs) are vesicles of 500–4000 nm in size that are generated from plasma membrane blebbing through apoptopodia formation or "beads-on-a-string" protrusions (Poon et al. 2014; Atkin-Smith et al. 2015). ABs sediment at 10,000 g and express phosphatidylserine (PS), annexin V, thrombospondin and C3b, and contain genomic DNA fragments, histones, and cell organelles (Fadok et al. 1998; van Engeland et al. 1998; Savill et al. 1992; Takizawa et al. 1996; Akers et al. 2013).

The most commonly found RBPs associated with EVs are themselves components of the miRNA processing machinery. For example, during exosome maturation, such as in late endosomes and MVB, GW182 interacts with Ago2 to protect loaded miRNAs from 3′–5′ exoribonuclease complex degradation (Wood et al. 2012). GW bodies are cytoplasmic foci where GW182 and Ago2 congregate with untranslated mRNAs, frequently co-localized with multivesicular bodies resulting in the preferential delivery of single-stranded mature miRNAs into exosomes (Gibbings et al. 2009). However, in pathological conditions such as cancer, double-stranded pre-miRNAs can also be delivered along with other RBPs. In tumor-derived microvesicles (TMVs), pre-miRNAs are complexed with the RISC components Dicer and Ago2 but also with Exportin-5 (Clancy et al. 2019). Exportin-

5 binds to double-stranded pre-miRNA in the nucleus, facilitating the crossing from the nuclear envelope to the cytoplasm and the trafficking of the pre-miRNA complex into shedding TMVs. In tumor-derived exosomes (TDEs), pre-miRNAs are associated with Ago2 and Dicer but also with another component of the RISC complex, TRBP (Melo et al. 2014).

The mechanisms of RBP–RNA complex formation within EVs remain largely unknown. In the case of exosomes, several different cell-specific mechanisms have been proposed. For example, one mechanism involves the ESCRT (Endosomal Sorting Complexes Required for Transport) complex and accessory protein, ALIX, that together bind to Ago2 and associated miRNAs, thereby favoring incorporation into exosomes (Iavello et al. 2016). An ESCRT-independent mechanism, involves the recruitment and activation of nSMAse2 within the lipid rafts of the MVB membrane, resulting in the transformation of sphingomyelin into phosphocholine and ceramide, leading to polymerisation of the ceramide molecules to create ceramide-rich platforms (CRPs), which can act as ILV budding sites (Trajkovic et al. 2008). Free RNA can be directly recruited to the EVs via its secondary structure, which can form a motif with an affinity for lipid rafts (Janas et al. 2006; Mańka et al. 2021) and the posttranscriptional addition of hydrophobic groups to RNA can increase potential interactions with these domains (Janas et al. 2012; Janas et al. 2015). In addition, RNAs can also bind to RBPs such as hnRNPA2B1, which have high affinity for CRPs (Villarroya-Beltri et al. 2013).

RBPs are not only responsible for the loading of RNA into exosomes, but can do so selectively in response to different stimuli. For example, inflammation causes caspase-1 to cleave RILP, which in turn mediates the interaction at the cell periphery of the ESCRT-0 component, Hrs, and the RBP FMR1 along with specific miRNAs containing the AAUGC motif, as well as the transport of the MVB (Wozniak et al. 2020; Progida et al. 2006; Wang and Hong 2006; Jordens et al. 2001). Upon serum starvation, an ESCRT-independent pathway, the LC3-Dependent EV Loading, and Secretion (LDELS) pathway, becomes activated. LC3 is present in the ILVs within the MVB and interacts with ATG7 (Autophagy Related 7) and ATG12 (Autophagy Related 12) proteins, to mediate conjugation with phosphatidylethanolamine (PE) once LC3 is cleaved, to generate LC3-II (Leidal and Huang 2020). LC3-II captures proteins within the MVB limiting membrane, incorporating them into ILVs for subsequent release as exosomes (Leidal and Huang 2020). Several RBPs can bind to LC3-II through the LIR (LC3 Interaction Region) including SAFB (Scaffold Attachment Factor B), hnRNPK, G3BP1 (GTPase Activating Protein (SH3 Domain) Binding Protein 1), LARP1 (La Ribonucleoprotein 1), and SF3A1 (Splicing Factor 3a Subunit 1) (Leidal and Huang 2020). LC3 indirectly stimulates nSMAse2 thereby promoting ILVs budding (Leidal and Huang 2020), in addition to regulating MVB acidification through the interaction with ATP6V1E1, a subunit of the V1V0-ATPase, controlling exosome release (Guo et al. 2017). However, RBPs do not always favor RNA loading into EVs but can also act to retain RNA within the donor cell. HnRNPU, for example, binds miRNA targets containing the AA/UGCU motif within the nucleus thereby reducing their cytoplasmic abundance and consequently their availability for vesicular export (Zietzer et al. 2020).

The sequence, secondary/tertiary structures, and posttranscriptional modifications of RNAs can also coordinate the loading into EVs. For example, miRNAs containing 5′-phosphorylated adenine or uracil residues are preferentially bound to Ago2 through their MID (Middle) domain (Frank et al. 2010), whereas the presence of a 3′ UUU motif favors interaction with the Lupus La protein and its packaging into exosomes (Temoche-Diaz and Shurtleff 2019). The EXO motif (GGAG) allows the binding of sumoylated hnRNPA2B1, which promotes miRNA loading into exosomes (Villarroya-Beltri et al. 2013; Wu et al. 2018), and similarly, the presence of a hEXO motif (GGCU) allows the interaction with the NURR (N-terminal Unit for RNA Recognition) domain of SYNCRIP (Synaptotagmin-binding Cytoplasmic RNA-Interacting Protein hnRNP-Q or NSAP1) (Santangelo et al. 2016). MiRNAs that can also interact with the RNA-binding region generated by the restructuration of the three RRM (RNA Recognition Motif) domains of SYNCRIP increase their binding affinity (Hobor et al. 2018). Another RBP, YBX1, can bind to specific RNA sequences through its cold-shock domain (CSD) or alternatively non-specifically, in a structural dependent manner, through its C-terminal domain (Kleene 2018). This protein can also co-localize with GW bodies regulating the sorting of tRNAs, snRNAs, H/ACA and C/D box snoRNAs, Y RNAs, and Vault RNAs into exosomes (Goodier et al. 2007; Gallois-Montbrun et al. 2007; Shurtleff et al. 2017).

Posttranscriptional modifications such as 3′-terminal 2′-*O*-methylation can increase the stability of miRNAs (and piRNAs) by avoiding 3′-5′ exoribonucleases degradation which can enhance the affinity for Ago2 (Lambowitz et al. 2007; Ohara et al. 2007; Liang et al. 2020). The addition of non-template nucleotides at the 3′ end of miRNAs can also alter their cell distribution, as it has been observed that miRNAs with 3′ adenylated ends, especially accumulate in the cytoplasm, whereas 3′ end uridylated miRNAs are enriched at the exosomes (Lambowitz et al. 2007).

2.3 Function of Circulating miRNAs

It remains unclear whether the release of circulating miRNAs from cells is selective or passive (or both). For example, Turchinovich et al. found that the majority of circulating miRNAs are not associated with EVs but free-floating, leading the authors to suggest that they are by-products of dead cells (Turchinovich et al. 2011). Consistent with this concept many other authors have reported that even when miRNAs are in EVs they reflect the expression of the host cells in a passive manner. For example, Zhou et al. found that the upregulation of *miR-105* in metastatic breast cancer cells was mirrored in exosomes (Zhou et al. 2014), as did Fang et al. who demonstrated increased levels of *miR-1247-3p* in both tissues and serum exosomes from hepatocellular carcinoma (HCC) patients (Fang et al. 2018). In contrast, many studies have shown that the composition of circulating miRNAs either free-floating or associated with EVs differs greatly from their respective donor cells suggesting selective active transport (Mittelbrunn et al. 2011). Indeed, some secreted miRNAs appear not to be present at all in their respective parental cells

(Valadi et al. 2007b). In HCC, for example, it was demonstrated that although *let-7* and *miR-21* are upregulated in serum-derived exosomes this is not the case for whole blood or associated tissues (Mjelle et al. 2019). Another study showed that of the 109 miRNAs identified as being upregulated in sera from renal cell carcinoma (RCC), only 36 of them were also upregulated in the tissue samples (Wulfken et al. 2011). Similar findings have been made in other biological fluids, including Pigati et al. who found that the bulk of *miR-451* and *miR-1246* produced by malignant mammary cells was released into blood, milk and ductal fluids, while the majority of miRNAs produced by nonmalignant cells were retained (Pigati et al. 2010).

Several authors have demonstrated that circulating miRNAs can function in cell–cell communication, although it remains unclear if this function is restricted only to miRNAs packaged into exosomes. In 2007, Valadi et al. first proposed extracellular miRNAs can act in a paracrine manner for intercellular communication (Valadi et al. 2007b). Consistent with this hypothesis circulating miRNAs have subsequently been demonstrated to be delivered to recipient cells where they can inhibit gene expression through the same mechanisms than cellular miRNAs (Kosaka et al. 2010b) including tumor growth and metastasis formation (Solé and Lawrie 2021).

The mechanism of how circulating miRNAs "decide" in which organ they will reside for the constitution of the metastatic niche is still unclear. In the case of EVs however, organ tropism appears to be determined by their size, charge, and the presence of specific molecules in the surface of their membranes. For example, ovarian cancer cells deliver exosomes specifically enriched in mannose and sialic acid glycoproteins (Escrevente et al. 2011). Glioblastoma cells release phosphatidylethanolamine-enriched exosomes while neuroblastoma cells release exosomes in which the presence or absence of CD63 determines targeting either to neural dendrites or to whole neurons and glial cells, respectively (Escrevente et al. 2011; Toda et al. 2015; Laulagnier et al. 2018). The presence of specific integrins, such as ITGαvβ5 can define the specific binding of exosomes to liver Kupffer cells, whereas the expression of ITGα6β4 and ITGα6β1 can target exosomes to either fibroblast or epithelial cells of the lung (Hoshino et al. 2015).

Once the exosomes are anchored to the surface of recipient cells, they are internalized by endocytosis (clathrin-mediated endocytosis, caveolin-dependent endocytosis, lipid-raft mediated endocytosis), micropinocytosis, phagocytosis or direct fusion of plasma membranes (Xu et al. 2013; Tian et al. 2014; Gurung et al. 2021), while vesicle-free miRNAs seem to be taken up by specific receptors in the cell surface (Chen et al. 2012). In the case of tumor-derived extracellular vesicles, heparan sulfate proteoglycans seem to be the favored internalizing receptor for lipid-raft mediated endocytosis (Christianson et al. 2013), which is increased upon hypoxia (Cerezo-Magaña et al. 2021).

In the next stage, encapsulated miRNAs are released and can modulate gene expression in the cytoplasm or, surprisingly, can also act as ligands for intracellular receptors. Both *miR-21* and *miR-29a* have been demonstrated to interact with TLR8 (Toll-Like Receptor 8) expressed at the surface of the endosomes of macrophages through their GU-rich motif, triggering a pro-metastatic inflammatory response and

promoting tumor growth and metastasis (Fabbri et al. 2012). Moreover, *miR-142* and *miR-223,* which are endogenously expressed in macrophages, are transferred to hepatocarcinoma cells through heterologous cell–cell contact and gap junctions to posttranscriptionally regulate target mRNA expression (Aucher et al. 2013).

2.4 Circulating ncRNAs as Cancer Biomarkers

2.4.1 MiRNAs

MiRNAs are attractive candidates as cancer biomarkers as their stability and expression profile can classify different types of tumors according to their diagnosis and stage much more robustly and accurately than mRNA expression profiles (Lu et al. 2005). The first report of circulating miRNAs as a diagnostic and prognostic biomarkers for cancer came in 2008 when it was observed that levels of *miR-155*, *miR-210*, and *miR-21* in the serum of the DLBCL patients were significantly different from healthy sera and associated with prognostic outcome (Lawrie et al. 2008). In the same year, Mitchell et al. reported circulating miRNAs in the plasma of PCa patients (Bartel 2004). Subsequently, the potential of circulating miRNAs has been expanded to many different cancer types, encompassing diagnostic, prognostic, and predictive biomarkers (van den Berg et al. 2020).

MiRNAs are ubiquitously dysregulated in cancer, both directly in tumor cells but also in the cells that constitute the tumor microenvironment (TME). These changes in miRNA levels are often reflected in biological fluids thereby offering great potential as non-invasive biomarkers for cancer patients. A great deal of publications reports the potential of circulating miRNAs as biomarkers to diagnose or classify tumors, impart prognostic information, or monitor tumor progression and/or response to the therapy. Some of the most important publications in this field are summarized in Table 2.1.

A number of publications have proposed that individual or signature miRNA levels have better diagnostic power than current diagnostic techniques. For example, in the sera of gastric cancer (GC) patients a 3-miRNA signature (*miR-18a, miR-181b*, and *miR-335*) was observed to have higher diagnostic accuracy with all patients' stages (AUC 0.86), including stage I patients (AUC: 0.85), than current blood markers CEA (Carcinoembryonic Antigen) and CA19-9 (Carbohydrate Antigen 19-9) (AUC: CEA 0.65; CA19-9 0.67) (Izumi et al. 2021). In another GC study by Yang et al., plasma-derived exosomal *miR-195-5p* and *miR-211-5p* levels could distinguish GC patients from healthy individuals with an AUC value of 0.820 compared to CEA (AUC: 0.541), CA19-9 (AUC: 0.622) and CEA + CA19-9 (AUC: 0.567) (Yang et al. 2021). Moreover, the authors found that expression levels of *miR-211-5p* correlated with lymph node metastasis, tumor stage, and overall survival (OS). In plasma from cervical cancer, expression levels of exosomal *let-7d-3a* and *miR-30d-5p* were decreased compared to healthy individuals and patients with benign stages with an AUC of 0.828, which was better than cytology (AUC:

Table 2.1 Deregulated microRNAs in cancer that is used as diagnostic, prognostic, and predictive biomarkers

Cancer	Biomark.	miRNA	Sample	Expr.	Technique	Cohort	AUC	Clinical associations	References
BCa	Pg	*miR-34a-5p*	Urine	Up	RT-qPCR	63 BCa vs. 37 HC	0.813	Short DFS	Cavallari et al. (2020)
		miR-193a-5p					0.775		
		miR-200a-3p					0.811	–	
CC	D	*Let-7d-3p*	Plasma (exo)	Down	Microarray RT-qPCR ddPCR	119 CC + CIN II-III vs. 84 HC + CIN I	0.88 0.83 (2-miRNA) 0.89 (2-miRNA + cytology)	–	Zheng et al. (2019)
		miR-30d-5p					0.79 0.83 (2-miRNA) 0.89 (2-miRNA + cytology)		
CRC	D/Pg	*miR-1290*	Serum	Up	Microarray RT-qPCR	211 CRC vs. 56 adenomas vs. 57 HC	0.718 (adenoma) 0.830 (CRC)	Recurrence	Imaoka et al. (2016)
	Pd	*let-7i*	Plasma (exo)	Down	Microarray RT-qPCR	205 CRC (84 CS vs. 74 CR + 47 pre-C)	6-miRNA panel 0.825	Resistance to FOLFOX or XELOX	Han et al. (2020)
		miR-16	Plasma (exo)	Down	Microarray RT-qPCR	208 CRC (84 CS vs. 74 CR + 47 pre-C)	6-miRNA panel 0.828	Resistance to FOLFOX or XELOX	Han et al. (2020)
		miR-17-5p	Plasma (exo)	Up	RT-qPCR	35 (16 PR vs. 19 non-R)	–	FOLFOX4 and FOLFIRI treatment response	Sur et al. (2021)
		miR-30e	Plasma (exo)	Down	Microarray RT-qPCR	209 CRC (84 CS vs. 74 CR + 47 pre-C)	6-miRNA panel 0.829	Resistance to FOLFOX or XELOX	Han et al. (2020)
		miR-92a	Plasma (exo)	Up	Microarray RT-qPCR	210 CRC (84 CS vs. 74 CR + 47 pre-C)	6-miRNA panel 0.830	Resistance to FOLFOX or XELOX	Han et al. (2020)

(continued)

Table 2.1 (continued)

Cancer	Biomark.	miRNA	Sample	Expr.	Technique	Cohort	AUC	Clinical associations	References
		miR-100	Plasma (exo)	Up	Microarray RT-qPCR	206 CRC (84 CS vs. 74 CR + 47 pre-C)	6-miRNA panel 0.826	Resistance to FOLFOX or XELOX; histological grade	Han et al. (2020)
		miR-144-5p	Plasma (exo)	Down	Microarray RT-qPCR	207 CRC (84 CS vs. 74 CR + 47 pre-C)	6-miRNA panel 0.827	Resistance to FOLFOX or XELOX	Han et al. (2020)
		miR-185-5p	Plasma (exo)	Up	RT-qPCR	35 (16 PR vs. 19 non-R)	–	FOLFIRI treatment response	Sur et al. (2021)
	Pg	*miR-21-5p*	Plasma	Up	RT-qPCR	103 CRC	0.675 (POM1) 0.715 (POM6) 0.715 (+CEA, POM1) 0.789 (+CEA POM6)	–	Fukada et al. (2021)
		miR-25-3p	Serum	Up	RT-qPCR	17 CRC vs. 12 HC	–	Metastasis	Wang et al. (2020a)
		miR-125b-5p	Plasma (exo)	Up	RT-qPCR	35 (16 PR vs. 19 non-R)	–	High tumor grade, stage	Sur et al. (2021)
		miR-130b-3p	Serum	Up	RT-qPCR	17 CRC vs. 12 HC	–	Metastasis	Wang et al. (2020a)
		miR-425-5p	Serum	Up	RT-qPCR	17 CRC vs. 12 HC	–	Metastasis	Wang et al. (2020a)
EC	D/Pg	*miR-15a-5p*	Plasma (Exo)	Up	Sequencing RT-qPCR ddPCR	140 EC vs. 118 HC	0.611 0.904 (+ CEA + CA125)	p53 status, muscular infiltration depth, large tumors	Zhou et al. (2021)
	D	*miR-106b-5p*						–	
		miR-107						–	

GC	D	*miR-18a*	Serum	Up	RT-qPCR	680 GC vs. 255 HC	0.87 (3-miRNA)	–	Izumi et al. (2021)
		miR-181b		Up	RT-qPCR			–	
		miR-335		Up	RT-qPCR			–	
	D/Pg	*miR-18a-3p*	Serum	Up	RT-qPCR	104 GC vs. 49 HC	–	Hp infection, tumor invasion	Tsai et al. (2020)
		miR-195-5p	Plasma (exo)	Up	RT-qPCR	108 GC vs. 108 HC	0.82 (2-miRNA)	–	Yang et al. (2021)
		miR-211-5p		Up	RT-qPCR			Lymph node metastasis, tumor stage and survival	
		miR-301a-3p	Serum (exo)	Up	RT-qPCR	20 GC vs. 10 HC	–	Peritoneal metastasis	Xia et al. (2020)
		miR-4286	Serum	Up	RT-qPCR	104 GC vs. 49 HC	–	Hp infection, tumor stage and size, lymph node metastasis	Tsai et al. (2020)
	Pg	*miR-15b-3p*	Serum	Up	RT-qPCR	30 GC vs. 30 HC	–	OS	Konishi et al. (2020)
	Pd	*miR-1229-3p*	Plasma	Up	Microarray RT-qPCR	80 GC (28 w vs.52 w/o recurrence)	0.807	Advanced T-stage, recurrences, peritoneal metastasis, RFS and OS	Nishibeppu et al. (2020)
Glioma	D/Pg	*miR-2276-5p*	Plasma (exo)	Down	RT-qPCR	124 glioma vs. 36 HC	0.8107	Survival	Sun et al. (2021)
	Pg	*miR-182*	Plasma	Up	RT-qPCR	112 glioma vs. 54 HC	–	KPS score, WHO grade, OS, DFS	Xiao et al. (2016)
		miR-210	Serum	Up	RT-qPCR	136 glioma vs. 50 HC	–	Tumor grade, prognosis	Lai et al. (2015)

(continued)

Table 2.1 (continued)

Cancer	Biomark.	miRNA	Sample	Expr.	Technique	Cohort	AUC	Clinical associations	References
Melanoma	Pd	*miR-1234-3p*	Plasma	Down	NGS	23 stage IV (13 non-R vs. 10 R)	–	CII response	Bustos et al. (2020)
		miR-4649-3p							
		miR-615-3p							
NSCLC	Pg/Pd	*miR-1246*	Serum	Up	RT-qPCR	59 NSCLC vs. 65 HD	–	Metastasis, tumor-size, EGFR TKI response	Wu et al. (2016)
		miR-1290						Tumor-size, EGFR TKI response	
	Pd	*miR-21-5p*	Plasma	–	sRNAseq RT-qPCR	50 (22 R vs. 28 non-R)	0.812–0.875 (2-miRNAs) 0.862–0.925 (3-miRNAs)	Response to anti-PD-1 or anti-PD-L1 therapy	Shukuya et al. (2020)
		miR-28-5p	Plasma	–	sRNAseq RT-qPCR	50 (22 R vs. 28 non-R)	0.812–0.875 (2-miRNAs) 0.862–0.925 (3-miRNAs)	Response to anti-PD-1 or anti-PD-L1 therapy	Shukuya et al. (2020)
		miR-125b-5p	Plasma	Up	Sequencing	9 NSCLC (5 PR vs. 4 PD) vs. 7 HD	–	PD-L1 immunotherapy efficacy	Peng et al. (2020)
		miR-199a-5p	Plasma	–	sRNAseq RT-qPCR	50 (22 R vs. 28 non-R)	0.812–0.875 (2-miRNAs) 0.862–0.925 (3-miRNAs)	Response to anti-PD-1 or anti-PD-L1 therapy	Shukuya et al. (2020)

		miR-200a	Plasma (EVs)	Up	Sequencing RT-PCR	10 NSCLC	–	EGFR-TKIs response	Lin (2021)
		miR-200c	Plasma (EVs)	Up	Sequencing RT-PCR	10 NSCLC	–	EGFR-TKIs response	Lin (2021)
		miR-200c-3p	Plasma	–	RNAseq RT-qPCR	50 (22 R vs. 28 non-R)	0.812–0.875 (2-miRNAs) 0.862–0.925 (3-miRNAs)	Response to anti-PD-1 or anti-PD-L1 therapy	Shukuya et al. (2020)
		miR-210	Plasma (EVs)	Down	Sequencing RT-PCR	10 NSCLC	–	EGFR-TKIs response	Lin (2021)
		miR-320b	Plasma	Up	Sequencing	9 NSCLC (5 PR vs. 4 PD) vs. 7 HD	–	PD-L1 immunotherapy efficacy	Peng et al. (2020)
		miR-320c							
		miR-320d							
		miR-758	Plasma (EVs)	Down	Sequencing RT-PCR	10 NSCLC	–	EGFR-TKIs response	Lin (2021)
PCa	D	*miR-24-3p*	Urine	Down	RT-qPCR	385 Pca vs. 42 BPH	0.86 0.92 (3-miRNAs)	–	Fredsøe et al. (2018)
		miR-30c-5p		Up			0.81 0.92 (3-miRNAs)		
		miR-222-3p		Down			0.83 0.92 (3-miRNAs)		

(continued)

Table 2.1 (continued)

Cancer	Biomark.	miRNA	Sample	Expr.	Technique	Cohort	AUC	Clinical associations	References
	Pg	*miR-26a-5p*	Urine (Exo)	Up	NGS RT-qPCR	41 BCR vs. 59 non-BCR	0.672	Survival after radical prostatectomy	Sharova et al. (2021)
		miR-99b-3p							
		miR-532-5p						Survival after radical prostatectomy, recurrence	

AUC, Area under the curve; BCa, Bladder cancer; BCR, Biochemical recurrence; BPH, Benign prostate Hyperplasia; Biomark., Biomarker; C, chemotherapy; CC, Cervical Cancer; CIN, cervical intraepithelial neoplasia; CR, chemoresistant; CRC, Colorectal cancer; CS, Chemosensitive; D, Diagnostic; ddPCR, Droplet Digital PCR; DFS, Disease-free survival; EC, Endometrial cancer; EVs, Extracellular vesicles; Exo, exosomes; Exp., Expression; GC, Gastric Cancer; HC, Healthy control; Hp, *Helicobacter pylori*; NGS, next generation sequencing; NSCLC, Non-small cell lung cancer; PCa, Prostate cancer; Pd, Predictive; PD, progressive disease; Pg, Prognostic; POM1, 1-month post-operation; POM6, 6-months post-operation; PR, Partial response; R, responders; RT-qPCR, quantitative reverse transcription PCR

0.766) (Zheng et al. 2019). Furthermore, the integration of the two miRNAs with cytological testing further increased the diagnostic power to 0.887. In a similar fashion, when the levels of plasma exosomal *miR-15a-5p* were combined with blood CEA and CA125 levels for endometrial cancer patients, an AUC value of 0.899 was obtained, higher than the individual components (Zhou et al. 2021). In another study, levels of three miRNAs (*miR-222-3p*miR-24-3p/miR-30c-5p*) derived from the urine of prostate cancer patients were found to have higher diagnostic accuracy than PSA (Prostate-Specific Antigen), with an AUC of 0.89–0.95 in all patients and an AUC of 0.89–0.97 for patients with PSA level $\leq$ 15 ng/mL (Fredsøe et al. 2018). In the same study, three different miRNAs (*miR-125b-5p*let-7a-5p/miR-151a-5p*) were associated with recurrence-free survival (RFS).

Circulating miRNA expression is often related to prognostic outcomes and the presence of metastasis in cancer patients. For example, in colorectal cancer (CRC), several miRNAs have been identified as independent prognostic factors including serum-derived *miR-1290* and plasma-derived *miR-21-5p* (Imaoka et al. 2016; Fukada et al. 2021), and *miR-21-5p* expression has also been associated with recurrence and progressive disease after surgical resection (Fukada et al. 2021). Wang et al. observed that levels of serum exosomal *miR-25-3p*, *miR-130b*, and *miR-425-5p* correlated with progression and metastasis in CRC patients and that in vitro, this miRNA promoted the M2 polarization of macrophages (Wang et al. 2020a). In GC, two circulating miRNAs (*miR-4286* and *miR-211-5p*) were found to correlate with the presence of lymph node metastasis (Tsai et al. 2020; Yang et al. 2021), while in another study *miR-301a-3p* and *miR-1229-3p* correlated with peritoneal metastasis (Xia et al. 2020; Nishibeppu et al. 2020). In addition, the levels of *miR-211-5p* (and *miR-15b-3p*) were also related to OS of patients (Yang et al. 2021; Konishi et al. 2020; Nishibeppu et al. 2020). In glioma, sera levels of *miR-210* and plasma levels of *miR-182* and *miR-2276-5p* have been demonstrated to correlate with OS (Sun et al. 2021; Xiao et al. 2016; Lai et al. 2015).

MiRNAs found in urine have also been demonstrated to have prognostic ability in bladder cancer, for example, the levels of *miR-34a*, *miR-200a*, and *miR-193a* could distinguish between high- and low-risk patients as well as being associated with shorter event-free survival (EFS) (Cavallari et al. 2020). In urine from PCa cancer patients, the expressions of three miRNAs (*miR-26a-5p*, *miR-532-5p*, and *miR-99b-3p*) were upregulated in exosomes from patients with biochemical recurrence (BCR), an event associated with an aggressive phenotype (Sharova et al. 2021).

Apart from diagnostic and prognostic biomarkers, circulating miRNAs have widely been used as predictive biomarkers for therapeutic response. In a study of plasma samples obtained from patients with non-small cell lung cancer (NSCLC), that were classified according to their response to anti-PD-1 or anti-PD-L1 immunotherapy, 32 miRNAs were identified from whole plasma and seven from plasma-derived EVs (Shukuya et al. 2020). Another study identified that plasma exosomal *miR-320d*, *miR-320c*, and *miR-320b* could predict the response to PD-1/PD-L1 inhibitors in advanced NSCLC (Peng et al. 2020). Also in NSCLC, 23 differentially expressed plasma exosomal miRNAs were identified between good responders and

poor responders in patients treated with EGFR-tyrosine kinase inhibitors (TKIs) (Wu et al. 2016; Lin 2021). In another study in late-stage NSCLC that received EGFR TKI treatment, levels of sera *miR-1246* and *miR-1290* were decreased in responders compared to non-responders (Wu et al. 2016). In advanced CRC, plasma-derived exosomal levels of *miR-17-5p* and *miR-185-5p* were found to be upregulated in patients that responded to FOLFOX (leucovorin, 5-fluorouracil, and oxaliplatin) or FOLFIRI (Leucovorin, 5-fluorouracil, and irinotecan) in conjunction with targeted therapies (Sur et al. 2021). In another study, a 6-miRNA signature (*miR-100, miR-92a*, *miR-16*, *miR-30e*, *miR-144-5p*, and *let-7i*) was identified in plasma-derived exosomes from CRC patients that could distinguish chemoresistance between patients treated with FOLFOX o XELOX (capecitabine plus oxaliplatin), with higher accuracy (AUC: 0.825) than traditional tumor biomarkers (Han et al. 2020). In plasma obtained from melanoma patients, a nine-miRNA signature was identified that could classify the response of patients with metastasis to anti-PD-1 or anti-PD1/anti-CTLA-4 immunotherapy (Bustos et al. 2020).

In addition to the expression of miRNAs themselves, posttranscriptional modifications of circulating miRNAs have also been proposed as potential cancer biomarkers. For example, methylation levels of miRNAs were observed to be increased in both pancreatic and CRC patient serum when compared to healthy donors (Cole et al. 2019). Furthermore, the authors found higher levels in the sera of pre-surgery than post-surgery patients, and increased levels of methylated *miR-17-5p* in the sera of pancreatic cancer patients even though it was expressed at low levels or was absent in controls.

Aside from the evaluation of individual miRNAs, the use of next generation sequencing (NGS) techniques allows the evaluation of the whole circulating transcriptome of cancer patients. For example, Xie et al. used the GEO (Gene Expression Omnibus) database to obtain miRNA expression patterns from osteosarcoma plasma samples ($n = 20$) and normal controls ($n = 15$) and found several dysregulated miRNAs, long non-coding RNAs (lncRNAs) and mRNAs (Xie et al. 2018). Bai et al. used NGS in the plasma of HCC patients to identify a miRNA expression profile distinct from healthy individuals including the downregulation of *miR-486-5p* in HCC patients (Bai et al. 2019). Another study identified 14 aberrantly expressed miRNAs by NGS from serum and tissue of patients with cervical cancer ($n = 133$) compared to healthy controls ($n = 106$) (Shukla et al. 2019). They found that *miR-17-5p*, *miR-32-5p* and *miR-454-3p* were upregulated while *miR-409-3p* was downregulated in cancer patients. Hallal et al. used NGS analysis on serum-derived EVs from glioblastoma (GBM) patients and identified that *miR-486-3p* and *miR-25-3p* were overexpressed and that expression was associated with tumor aggressiveness (Hallal et al. 2020).

2.4.2 LncRNAs and Other ncRNAs

LncRNAs are transcripts of more than 200 nt in length that do not encode for a protein. Although initially considered as transcriptional by-products, lncRNAs show higher conservation than either introns or intergenic sequences, particularly in their secondary and tertiary structures, suggesting a functionality maintained across evolution (Guttman et al. 2009; Ponjavic et al. 2007; Tavares et al. 2019; Smith et al. 2013). Moreover, lncRNA genes share many characteristics of mRNAs such as histone modifications (i.e., H3K4me3, H3K36me3, H3K27Ac, and H3K4me1) (Djebali et al. 2012; Ayupe et al. 2015), 5′ capping, 3′ adenylation and transcript splicing (Ayupe et al. 2015; Melé et al. 2017). Interestingly, lncRNAs exhibit a significantly higher turnover rate than mRNAs, along with more tissue- and tumor-specific expression patterns, making them attractive candidates as biomarkers (Clark et al. 2012; Cabili et al. 2011; Iyer et al. 2015; Geisler and Coller 2013).

Several circulating lncRNAs have been demonstrated to be potentially useful cancer biomarkers. Indeed, the lncRNA, urinary *PCA3* has FDA approval for prostate cancer diagnosis (Marks et al. 2007). In GC patients, for example, serum levels of *HOTTIP* (AUC: 0.827) and plasma levels of *lncRNA-GC1* (AUC: 0.903) were observed to have better discriminatory capability than either CEA (AUC: 0.653 in serum and 0.5987 in plasma), CA72–4 (AUC: 0.639 in serum and 0.6816 in plasma), or CA19–9 (AUC: 0.685 in serum and 0.6482 in plasma) (Zhao et al. 2018a; Guo et al. 2020). In addition, *HOTTIP* upregulation was observed to associate with poor prognosis indicators such as advanced TNM stage, presence of invasion, and decreased OS. Another lncRNA, *LncUEGC1* (AUC: 0.8406) was even able to distinguish patients at early stages of GC from those with an underlying chronic inflammation disease, with a better accuracy than CEA (AUC: 0.6123) (Lin et al. 2018). In CRC patients, serum levels of *CRNDE* (AUC: 0.892) and plasma levels of *RPPH1* (AUC: 0.860) show a better diagnostic value than CEA (AUC: 0.688 in serum and 0.790 in plasma) (Liu et al. 2016; Liang et al. 2019). Moreover, *CRNDE* upregulation also associates with clinical parameters like the presence of metastasis at lymph nodes and distal organs. In plasma of esophageal squamous cell carcinoma (ESCC) patients, *POUF3* (AUC: 0.842) levels had a higher degree of diagnostic accuracy than the tumor marker SCCA (Squamous Cell Carcinoma Antigen) (AUC: 0.784) (Tong et al. 2015). Additionally, it has also recently been described that lncRNA fragments, like the S fragment of *RN7SL1* (AUC: 0.870) contained within plasma exosomes of HCC patients, can be used as circulating biomarkers with better discriminatory ability than their respective full-length transcriptional counterparts (AUC: 0.750) (Tan et al. 2019).

LncRNAs can also exhibit predictive biomarker capabilities. For example, increased levels of serum *PART1* were associated with poor response to gefitinib treatment in ESCC patients (Kang et al. 2018), and in the serum of GBM patients, high levels *of lncSBF2-AS1* were demonstrated to be associated with resistance to temozolomide treatment (Zhang et al. 2019). The upregulation of *UCA1* in serum of ovarian cancer patients was found to be indicative of cisplatin resistance (Li et al.

2019a). In addition, increased levels of *lncARSR* in plasma of RCC patients were observed to be associated with response to sunitinib treatment (Qu et al. 2016a).

Apart from lncRNAs and miRNAs, other circulating small non-coding molecules such as small nucleolar RNAs (snoRNAs), small nuclear RNAs (snRNAs), piwi-associated RNAs (piRNAs), transfer RNAs (tRNA), and circular RNAs (circRNA) have also been reported in body fluids (Table 2.2) (Umu and Langseth 2018). For example, Kitagawa et al. found that levels of snoRNAs are upregulated in sera of renal and pancreatic cancer patients compared to healthy controls (Zhao et al. 2020a; Kitagawa et al. 2019) and that *SNORA74A* and *SNORA25* could distinguish between healthy and early-stage pancreatic cancer patients with an AUC $>$ 0.9 (Kitagawa et al. 2019). In NSCLC patients, levels of the snoRNAs *SNORD33*, *SNORD66*, *SNORD76*, *SNORD78*, or *SNORA42* from either plasma or sputum could distinguish between patients with chronic obstructive pulmonary disease, smokers, or healthy individuals (Liao et al. 2010; Su et al. 2016). The snRNA *RNU2-1f* has been found to be highly expressed in sera and plasma from colorectal, pancreatic, lung, and ovarian cancer patients (Richter et al. 2014; Baraniskin et al. 2013; Köhler et al. 2016; Kuhlmann et al. 2014). Another study observed that the snRNA *RNU6–1* was overexpressed in serum exosomes of GBM patients compared with healthy controls and was associated with the differentiation of GBM from nonneoplastic brain lesions and primary central nervous system lymphoma (Puigdelloses et al. 2020). Circulating piRNAs (i.e., *piR-001311*, *piR-004153*, *piR-017723*, *piR-017724*, and *piR-020365*) have been identified in the sera of CRC patients and found to be downregulated compared to healthy individual samples, with *piR-017724* demonstrated to be an independent prognostic biomarker (Gállego Pérez-Larraya et al. 2019). Another study also in CRC sera, observed an upregulation of *piR-5937* and *piR-28876* to be associated with tumor stage postoperative status (Vychytilova-Faltejskova et al. 2018). Circulating tRNA-derived fragments (tRFs) (i.e., *tRF-Glu-CTC-003, tRF-Gly-CCC-007, tRF-Gly-CCC-008, tRF-Leu-CAA-003, tRF-Ser-TGA-001,* and *tRF-Ser-TGA-002*) were identified in plasma of breast cancer patients and were found to be significantly downregulated, being their expression levels associated with prognostic outcome (Wang et al. 2020b). In GC patients, the levels of *hsa-tsr016141* and *tiRNA-5034-GluTTC-2* in serum and plasma, respectively, were associated with prognosis (Gu et al. 2021; Zhu et al. 2019).

CircRNAs, similar to miRNAs, are highly resistant to exonuclease-mediated degradation and consequently are widely believed to be one of the most stable RNA types present in biofluids (Yu et al. 2019). Several circRNAs have been identified in biological fluids (Table 2.2) including *circSMARCA5* and *circ_0000190* expression that has shown to correlate with AFP and CA19-9 levels, clinical markers for HCC, and GC diagnosis, respectively (Chen et al. 2017; Li et al. 2019b).

Table 2.2 Noncoding RNAs deregulated in biofluids of cancer patients are used as diagnostic, prognostic, and predictive biomarkers

	Cancer	Biomark.	ncRNA	Sample	Expr.	Technique	Cohort	AUC	Clinical associations	References
LncRNAs	CRC	D	*RPPH1*	Plasma (Exo)	Up	RT-qPCR	52 CRC vs. 41 HC	0.860	Surgery response	Liang et al. (2019)
		D/Pg	*CRNDE*	Serum (Exo)	Up	RT-qPCR	148 CRC vs. 240 CBD vs. 80 HC	0.892	Lymph node and distant metastasis	Liu et al. (2016)
	ESCC	D	*POU3F3*	Plasma	Up	RT-qPCR	147 ESCC vs. 123 HC	0.842	–	Tong et al. (2015)
		Pd	*PART1*	Serum (Exo)	Up	RT-qPCR	42 R vs. 37 non-R	0.839	Gefitinib response	Kang et al. (2018)
	GBM	Pd	*LncSBF2-AS1*	Serum (Exo)	Up	RT-qPCR	10 recurrent vs. 10 primary	–	Temozolomide response	Zhang et al. (2019)
	GC	D	*lncRNA-GC1*	Plasma (Exo)	Up	RT-qPCR	522 GC vs. 85 GPL vs. 219 HC	0.9033	Early stage	Guo et al. (2020)
			lncUEGC1	Plasma (Exo)	Up	RT-qPCR	51 GC vs. 18 CAG vs. 60 HC	0.8406	Early stage	Lin et al. (2018)
		D/Pg	*HOTTIP*	Serum (Exo)	Up	RT-qPCR	126 GC vs. 120 HC	0.827	TNM, invasion, OS	Zhao et al. (2018a)
	HCC	D/Pg	*RN7SL1 (S fragment)*	Plasma (Exo)	Up	RNA-Seq RT-qPCR	100 HCC vs. 20 CHB vs. 30 HC	0.782 (stage A) 0.927 (stage B) 0.924 (stage C) 0.869 (all)	Survival	Tan et al. (2019)
	OC	Pd	*UCA1*	Serum (Exo)	Up	RT-qPCR	24 R vs. 32 non-R	–	Cisplatin response	Li et al. (2019a)
	PCa	D	*PCA3*	Urine	Up	RT-qPCR	60 PCa vs. 166 BPH	0.678	–	Marks et al. (2007)
	RCC	Pd	*LncARSR*	Plasma	Up	RT-qPCR	38 R vs. 33 non-R	–	Sunitinib response	Qu et al. (2016a)
SnoRNAs	NSCLC	D	*SNORD33*	Plasma	Up	RT-qPCR	37 NSCLC vs. 26 COPD vs. 22 HC	0.8233 (HC) 0.82 (COPD)	Discriminate from COPD and smokers (Sm)	Liao et al. (2010)
				Sputum	Up			0.7230 (Sm)		
			SNORD66	Plasma	Up			0.8139 (HC) 0.7903 (COPD)		

(continued)

Table 2.2 (continued)

	Cancer	Biomark.	ncRNA	Sample	Expr.	Technique	Cohort	AUC	Clinical associations	References
				Sputum				0.8065 (Sm)		
			SNORD76	Plasma	Up	RT-qPCR	67 NSCLC vs. 69 smokers	0.8064 (HC) 0.8149 (COPD)		Su et al. (2016)
			SNORD78	Sputum	Up			0.8112		
			SNORA42		Up			0.7431		
	PDAC	D	*SNORA25*	Serum	Up	RT-qPCR	27 PDAC vs. 13 BGD	0.903 (HC) 0.914 (0/I/IIA) 0.903 (IIB/III/IV)	Early diagnosis	Kitagawa et al. (2019)
			SNORA74A	Serum	Up	RT-qPCR	27 PDAC vs. 13 BGD	0.909 (HC) 0.952 (0/I/IIA) 0.909 (IIB/III/IV)	Early diagnosis	
SnRNAs	CRC	D	*RNU2-1f*	Serum/plasma	Up	RT-qPCR	132 CRC vs. 20 colon adenomas vs. 129 HC	0.972	Tumor stage, surgery response	Baraniskin et al. (2013)
	LC	D/Pd	*RNU2-1f*	Serum	Up	RT-qPCR	211 LC vs. 56 chronic lung disease vs. 58 HC	0.910	MS in IIIB/IV stages, chemotherapy response	Köhler et al. (2016)
	OC	D/Pd	*RNU2-1f*	Serum	Up	Microarray RT-qPCR	63 before vs. 56 after AC vs. 35 HC	0.740	Recurrence, platinum chemotherapy response	Kuhlmann et al. (2014)
	GBM	D	*RNU6–1*	Serum	Up	ddPCR	18 GBM vs. 30 stroke vs. 30 hemorrhage vs. 18 acute demyelinating lesions vs. 21 brain metastases vs. 12 PCNSL vs. 30 HC	0.759 (HC) 0.695 (Stroke) 0.724 (Hemo) 0.728 (AMSP) 0.700 (All non-tumor lesions)	Differentiation GBM from non-neoplastic brain lesions and PCNSL	Puigdelloses et al. (2020)

piRNAs	CRC	D	*piR-5937*	Serum	Down	Small RNA-Seq RT-qPCR	403 CRC vs. 276 HC	0.7567	Tumor stage	Vychytilova-Faltejskova et al. (2018)
			piR-28,876		Down			0.7869		
			piR-001311	Serum	Down	HTS Seq RT-qPCR	240 CRC vs. 230 HC	0.867 (5 piRNAs)	–	Qu et al. (2019)
			piR-02036		Down					
			piR-004153		Down					
			piR-017723		Down					
		Pg	*piR-017724*	Serum	Down	HTS Seq RT-qPCR	240 CRC vs. 230 HC	0.867 (5 piRNAs)	OS, PFS	Qu et al. (2019)
tRNAs	BC	D/Pg	*tRF-Ser-TGA-001*	Plasma	Down	RNA-Seq RT-qPCR	176 early-stage BCa vs. 140 HC	0.844	BCa subtype, invasion and lymph node metastasis	Wang et al. (2020b)
			tRF-Ser-TGA-002		Down					
			tRF-Leu-CAA-003		Down					
			tRF-Glu-CTC-003		Down					
			tRF-Gly-CCC-007		Down					
			tRF-Gly-CCC-008		Down					
	GC	D/Pg	*hsa_tsr016141*	Serum	Up	RT-qPCR	130 GC vs. 50 gastritis vs. 63 postoperative GC vs. 110 HC	0.814 (HC) 0.692 (Gastritis)	Tumor grade, lymph node metastasis, surgery response	Gu et al. (2021)
		Pg	*tiRNA-5034-GluTTC-2*	Plasma	Down	RT-qPCR	86 GC vs. 37 HC	0.835	Tumor size, OS	Zhu et al. (2019)
circRNAs	GC	D/Pg	*circ_0000190*	Plasma	Down	RT-qPCR	104 GC vs. 104 HC	0.600	Tumor size, TNM, lymph node, and distant metastasis, CA19–9 levels	Chen et al. (2017)
	HCC	Dg/Pg	*circSMARCA5*	Plasma	Down	RT-qPCR	135 HCC vs. 143 cirrhotic vs. 117 hepatitis vs. 103 HC	0.711 (Cirrhotic) 0.853 (Hepatitis) 0.938 (HC)	Tumor size, differentiation, TNM, invasion, AFP levels	Li et al. (2019b)

(continued)

Table 2.2 (continued)

	Cancer	Biomark.	ncRNA	Sample	Expr.	Technique	Cohort	AUC	Clinical associations	References
	OSCC	D	*circ_0001874*	Saliva	Up	RT-qPCR	93 OSCC vs. 85 HC	0.863	TNM and tumor grade	Zhao et al. (2018b)
			circ_0001971		Up	RT-qPCR		0.845	TNM	

AC, Adjuvant Chemotherapy; ADC,Adenocarcinoma; AUC, Area under the curve; AMPS, Acute Multiple Sclerosis Plaques; BC, Breast cancer; BCa, Bladder cancer; BCT, Benign Colorectal Tumor; BGD, Benign Gastric Disease; Biomark., Biomarker; BPH, Benign prostate Hyperplasia; circRNA, Circular RNA; CAG, Chronic Atrophic Gastritis; CDB: Colorectal Benign Disease; CHB, Chronic Hepatitis B; COPD, Chronic Obstructive Pulmonary Disease; CRC, Colorectal cancer; D, Diagnostic; ddPCR, Droplet Digital PCR; DgC, Digestive Cancer; ESCC, Esophageal squamous cell carcinoma; Exo, exosomes; Exp., Expression; GBM, Glioblastoma multiforme; GC, Gastric Cancer; GPL, Gastric Precancerous Lesions; HC, Healthy control; HCC, Hepatocellular carcinoma; LC, Lung cancer; lncRNA, Long non-coding RNA; MS, Median Survival; ncRNA, non-coding RNA; NSCLC, Non-small cell lung cancer; OC, Ovarian cancer; OS, Overall Survival; PCa, Prostate cancer; PCNSL, Primary Central Nervous System Lymphoma; Pd, Predictive; PDAC, Pancreatic ductal adenocarcinoma; PFS, Progression-free survival; Pg, Prognostic; piRNA, piwi-interacting RNA; R, Responders; RCC, Renal cell carcinoma; RFS, Relapse-free survival; RT-qPCR, quantitative reverse transcription PCR; SCC, Squamous cell carcinoma; Sm, Smoker; snoRNAs, Small nucleolar RNAs; snRNAs, Small nuclear RNAs; tRNAs, Transfer RNA; US, Urinary Sedimentation

2.5 Methods and Challenges to Detect Circulating miRNAs

Currently, the gold standard for diagnostic and other biomarker detection relies upon the invasively obtained tissue biopsy. However, this approach is expensive and not without risk to the patient. Although circulating miRNAs (and other RNAs) hold great clinical potential as noninvasive biomarkers, the translation of this knowledge into routine clinical practice still leaves a lot to be desired. The studies mentioned above are nearly all single-center retrospective studies and generally the cohorts in these studies are insufficiently powered (Tables 2.1 and 2.2). Consequently, there are many nonoverlapping and even contradictory circulating miRNA markers resulting primarily from biological and technical variations caused by differences in (1) starting material including differences in techniques of purification of EVs, between whole plasma and sera samples, control populations used, RNA extraction methods, etc.; (2) detection technology (e.g., microarray, qRT-PCR, and NGS) (Fig. 2.2); (3) altered signal/noise ratio of target miRNAs due to the presence of secondary response-associated miRNAs (e.g., inflammatory and immune); (4) use of differing statistical models for discovery and validation. In addition to these differences, the general low quality and low quantity of circulating miRNAs in biological fluids further compound the problem.

The choice of biological fluid itself can create challenges. For example, in some biological fluids such as urine, saliva, sputum, or stool samples, bacterial contamination can be particularly problematic. Even in the case of more "sterile" samples such as serum and plasma, the expression levels of the RNAs can differ greatly between blood fractions due to hemolysis. Moreover, the inter-patient variability in these samples is increased because of the release of EV-associated RNAs from stimulated platelets (McDonald et al. 2011; Kirschner et al. 2011; Pritchard et al. 2012). Also, intra-patient variability between RNA levels depends on the time of the day, diet, gender, age, alcohol consumption, etc. (Heneghan et al. 2010; Gourzones et al. 2013; Hu et al. 2013).

2.6 Bioinformatics Analyses

Once dysregulated miRNAs have been identified by detection techniques bioinformatics tools are frequently used to identify putative target genes. Such tools rely on a combination of experimentally validated miRNA target genes and various predictive algorithms.

Predictive algorithms generally fall into one of the following categories or a combination thereof: thermodynamic (free energy hybridization of the RNA hybrid), evolutionary (sequence conservation), sequence-based (such as seed match, site accessibility, 3′ compensatory pairing, G:U pairs allowed in the seed or local AU content), or probabilistic (target: site abundance), to which they apply either Bayesian Networks or other machine learning methods. The most well-known prediction

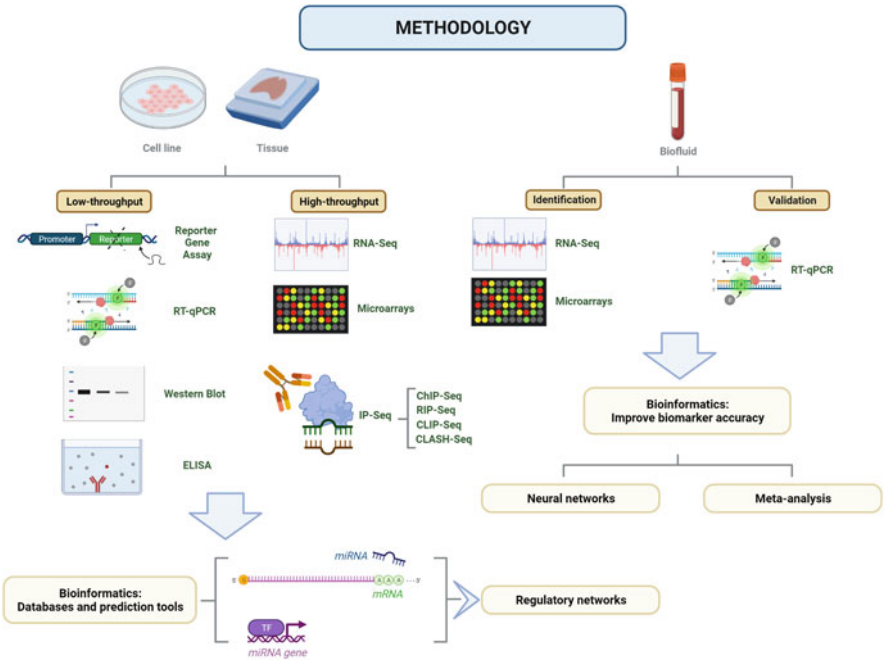


Fig. 2.2 Methods for microRNA detection. The transcription factors (TFs) that regulate miRNA transcription and the miRNA targets are experimentally identified in cell lines and tissues of different origins either by low-throughput techniques (reporter gene assays, RT-qPCR, Western Blot, or ELISA) or by high-throughput techniques (microarrays, RNA-Seq, and IP-Seq). The results obtained can be collected in databases that serve for the construction of software that predict TF:miRNA and miRNA:RNA interactions to establish regulatory networks. Some of these miRNAs can also be identified in human biofluids by RNA-Seq and microarray and further validated using RT-qPCR techniques and the evaluation of their biomarker properties can be improved when applying neural networks or meta-analysis

algorithms are miRanda, PicTar, PITA, TargetScan, StarMir, SVmicrO, DianamicroT, and mirTarget (John et al. 2004; Krek et al. 2005; Kertesz et al. 2007; Agarwal et al. 2015; Liu et al. 2010; Rennie et al. 2014; Paraskevopoulou et al. 2013; Liu and Wang 2019). An increasing number of databases also integrate expression levels of miRNAs and target transcripts, obtained from deep-sequencing analysis, as an additional parameter for prediction. Other features can also include the biophysical constraints forced by Argonaute protein on miRNA:target interaction or the presence of miRNA seed and seedless target sites at the CDS and 5′ UTR (Khorshid et al. 2013; Hsu et al. 2011; Jeggari et al. 2012; Rennie et al. 2016). Some newer predictive algorithms such as TargetMiner and MultimiTar are based on Support Vector Machine (SVM) classifiers that include negative examples of miRNA targets in order to increase the robustness of the predictions (Bandyopadhyay and Mitra 2009; Mitra and Bandyopadhyay 2011).

2.6.1 Databases of miRNA: Target RNA Interaction

Prediction algorithms are typically used to generate databases and/or trained using existing databases that collect experimentally validated miRNA targets (Fig. 2.2). For example, DIANA-TarBase is a database that compiles information about experimentally validated miRNA targets in 600 different cell lines and tissue types in more than 450 experimental conditions. The methodologies comprise low-throughput experiments like reporter gene assays, qPCR, Western Blot, or ELISA, as well as high-throughput experiments like microarrays, RNA-Seq, RIP-Seq, RPF-Seq, CLIP-Seq, AGO CLIP-seq, CLEAR-CLIP, or CLASH obtained from DNA Data Bank of Japan (DDBJ) and Gene Expression Omnibus (GEO) repositories (Karagkouni et al. 2018). Similarly, developers of the starBase database analyzed CLIP-Seq data from GEO, but considered all the coding and non-coding competing endogenous RNAs (especially lncRNAs, pseudogenes and circRNAs) with which a miRNA can interact (Li et al. 2014a). MiRGator compiles the information of deep-sequencing results of public databases such as GEO, SRA (Short Read Archive), and TCGA (The Cancer Genome Atlas). These datasets were manually curated into different tissue and disease types. During the alignment process, mismatches on miRNA loci are permitted, thus allowing the identification of iso-miRs and miRNA variants with 3′ non-templated nucleotide additions (Cho et al. 2013). The PolymiRTS database contains both DNA polymorphisms and INDELs at miRNA seed regions and at the 3′ UTR of their target sites. The SNPs and INDELs of mature miRNAs were obtained from comparisons with the dbSNP and miRBase database repositories; whereas target gene SNPs and INDELs were retrieved from the dbSNP database and mapped onto experimentally validated CLASH sequences (Bhattacharya et al. 2014). Other databases, such as miRTarBase and miRSel (Huang et al. 2020; Naeem et al. 2010), have been created by text mining articles from the PubMed repository for miRNA–gene associations.

2.6.2 Databases of Transcription Factor: miRNA Gene Interaction

Other databases focus on the study of the upstream regulation of miRNA transcription by transcription factors (TFs). For example, the TransmiR database uses manually curated publications of TF-miRNA regulation supported by the presence of highly conserved TF binding motifs at the promoters of miRNAs or by luciferase reporter and ChIP-seq experiments. (Tong et al. 2019). The MirTrans database includes cell type-specific information on miRNA promoters, transcription factor binding sites, and transcriptional start sites (TSSs), identified by the presence of H3K4me3 marks and DNAse-I hypersensitive sites. Moreover, the transcriptional regulation of miRNAs is experimentally supported by GRO/PRO-Seq and CAGE (Cap Analysis of Gene Expression) techniques, among others (Hua et al. 2018). The

DIANA-mirGen database uses machine learning algorithms trained on more than 1000 CAGE experiments corresponding to 133 different tissues, cell lines and primary cells available in the FANTOM repository, allowing cell-specific TSS annotation. In combination with DNA-Seq and ChIP-Seq datasets from ENCODE, DIANA-mirGen can identify the DNAse hypersensitivity sites containing the transcription factor motifs within the defined TSS (Perdikopanis et al. 2021).

ChIPBase includes the results of more than 500 ChIP-Seq experiments carried out on different cell types and tissues from six different model organisms, allowing the genome-wide identification of the binding sites of 252 transcription factors and their positions relative to the TSSs of miRNAs and other ncRNAs (Zhou et al. 2017). The TSmir database generates expression patterns of tissue-specific miRNAs and the transcription factors that regulate them, with a special focus on transcription factors expressed in the same tissue creating tissue-specific regulatory networks (Guo et al. 2014).

The CircuitsDB database contains information on regulatory loops between transcription factors and miRNAs that commonly regulate genes, in other words, mixed miRNA/TF Feed-Forward regulatory Loops (FFLs) and Feed-Back regulatory Loops (FBLs). As the classification of feedback loops as positive or negative is tissue-dependent, the expression values across 14 human tissues of the transcription factor are also included (Friard et al. 2010). Similarly, TMREC explores linear regulatory cascades (transcription factor-miRNA-mRNA), following the Breadth First Search (BFS) traversal method in the context of human diseases, including cancer (Wang et al. 2015).

2.6.3 Construction of Regulatory Networks

Both miRNA: target RNA and TF:miRNA gene interaction databases can be used to reconstruct regulatory networks to infer their possible biological impact. Some of the above-mentioned databases, for example, starBase, miRGator, PolymiRTS, or CircuitsDB, can include a module of functional analyses using KEGG (Kyoto Encyclopedia of Genes and Genomes), DAVID (Database for Annotation, Visualization, and Integrated Discovery), GO (Gene Ontology), BioCarta, Reactome or PID (Pathway Interaction Database) databases (Ogata et al. 1999; Dennis et al. 2003; Harris et al. 2004; Vastrik et al. 2007; Schaefer et al. 2009). Additionally, those databases focus on the study of SNPs or somatic mutations from genome-wide association studies (GWAS) using data from the NHGRI GWAS Catalog, dbGaP, or Cancer GAMAdb to predict the effect of these genetic variations on the risk of developing a disease, including cancer (Buniello et al. 2019; Mailman et al. 2007; Schully et al. 2011).

2.6.4 Databases of miRNAs with Cancer Biomarker Potential

In addition to the databases listed above, there are also disease-focused databases such as the HMDD (Human MicroRNA Disease Database), Mir2Disease, miRCancer, TUMIR, or OncomiRBD (Li et al. 2014b; Jiang et al. 2009; Xie et al. 2013; Dong et al. 2013; Wang et al. 2014). The dbDEMC database collates miRNA expression profiles from different cancer types obtained from the GEO and TCGA databases, microarray, and NGS data. SomamiR is a database of cancer-associated somatic mutations found in miRNAs and seed regions of target genes and also lncRNAs and circRNAs identified by CLASH, PAR-CLIP, and HITS-CLIP experiments or predicted by mapping somatic mutations from COSMIC to the 3′ UTRs of coding genes (from RefSeq annotations) and non-coding genes (from LNCipedia and circBase annotations) (Bhattacharya et al. 2013; Bhattacharya and Cui 2016).

More recently, databases have been created specifically dedicated to extracellular circulating ncRNAs, including miRNAs. For example, miRandola is a manually curated database based on literature mining (Russo et al. 2012). The ExRNA Atlas and CMEP (Circulating MicroRNA Expression Profiling) databases contain expression profiles of extracellular small non-coding RNAs in different human biofluids derived from RNA sequencing, miRNA microarray, and qRT-PCR techniques (Murillo et al. 2019; Li et al. 2019c).

2.6.5 Application of Neural Networks and Meta-Analysis to Identify Robust Biomarkers

The construction of neural network (NN) models using miRNA profiles in human fluids can be used to create a formula for identifying cancer. For example, in colon cancer, *hsa-miR-6726-5p, hsa-miR-7111-5p, hsa-miR-1247-3p*, and *hsa-miR-614* expression levels in serum were used for designing a NN model that can diagnose colon cancer with an AUC = 1 (Afshar et al. 2019). In epithelial ovarian cancer (EOC), the expression levels of a signature of seven microRNAs (*miR-29a-3p, miR-92a-3p, miR-200c-3p, miR-320c, miR-335–5 p, miR-450b-5p,* and *miR-1307–5 p*) in serum was used to create a NN EOC diagnosis model with an AUC = 0.89 (Elias and Fendler 2017). The expression values of *miR-15a*, *miR-101*, and *miR-144* in the plasma of women with BI-RADS (Breast Imaging Reporting and Data System) were used to create a NN model to discern whether the tumor is benign or malignant with an AUC = 0.96 (Pezuk et al. 2017). Although the use of NN models to increase the accuracy of cancer diagnosis using circulating miRNAs holds great promise it is still in its infancy. Instead, many studies have chosen a meta-analysis approach to generate meaningful data models (Table 2.3). The most analyzed miRNA in these analyses is *miR-21* (Peng et al. 2017; Gao et al. 2016a, b; Jinling et al. 2017; Qu et al. 2016b). For example in CRC *miR-21* was found to have a higher overall diagnostic power in serum samples than plasma (Peng

Table 2.3 Meta-analysis studies of miRNAs in noninvasive sources of biological material as cancer biomarkers

Cancer	Biomark.	Sample origin	N° miRNAs	Studies	Subjects	Sensitivity	Specificity	AUC	References
BC	D	Plasma/serum/tissue	1 (*miR-21*)	11	1531	0.72	0.8	0.85	Gao et al. (2016b)
			1 (*miR-34a*)	13	2352	0.855	0.7	0.8	Imani et al. (2017)
	Pg	Serum/BM/tissue	1 (*miR-21*)	7	1629	–	–	–	Jinling et al. (2017)
		Blood	1 (*miR-203*)	6	413	–	–	–	Shao et al. (2017)
BCa	Pg	Serum/urine/tissue	6	26	2753	–	–	–	Forberg et al. (2018)
Colonic adenoma	D	Fecal	6	10	946	0.573	0.761	0.747	Yau et al. (2019)
CRC			21	46	3316	0.588	0.848	0.811	
		Plasma/serum	4	19	6010	0.833	0.931	0.943	Carter et al. (2017)
		Serum	1 (*miR-21*)	57	2214	0.75	0.84	0.87	Peng et al. (2017)
	Pg	Blood	1 (*miR-203*)	6	330	–	–	–	Shao et al. (2017)
		Serum/BM/tissue	2	26	2653	–	–	–	Pop-Bica et al. (2018)
Digestive	D	Plasma/serum/tissue	1 (*miR-139*)	4	405	0.89	0.91	0.96	Wang et al. (2018)
DLBCL	Pd	Serum	8	3	230	–	–	–	Ting et al. (2019)
Gastrointestinal	Pg	Plasma/serum/blood/ BM/tissue	59	60	6225	–	–	–	Zheng et al. (2017)
GC	D	Plasma	1 (*miR-106*)	5	585	0.77	0.82	0.83	Peng et al. (2018)
		Serum				0.64	0.82	0.73	
	Pg		1 (*miR-130a*)	12	2141	–	–	–	Peng and Duan (2019)
Glioma	D	CSF	1 (*miR-21*)	81	7494	0.88	0.89	0.94	Yin et al. (2016)
		Plasma/serum/CSF	1 (*miR-21*)	4	206	0.86	0.94	0.88	Zhou et al. (2018)
		Plasma/serum/CSF/ tissue	Single miRNA	34	3605	0.83	0.85	0.9	
		Blood/plasma/serum/ CSF/tissue	miRNA panels	17	2032	0.9	0.95	0.97	
				51	5091	0.85	0.9	0.93	

		Blood/plasma/serum/CSF/tissue	Overall miRNAs						
		Blood/plasma/serum				0.84	0.85	0.92	
		CSF/tissue				0.89	0.98	0.98	
HCC	D	Plasma/serum	21 miRNAs	50	7713	0.758	0.75	0.82	Huang et al. (2016)
			1 (*miR-21*)			0.866	0.795	0.88	
			1 (*miR-122*)			0.68	0.733	0.77	
	Pg	Serum	1 (*miR-130b*)	15	3300	–	–	–	Peng and Duan (2019)
HNSCC	D	Saliva	23	17	759	0.697	0.868	0.803	Kang and Eun (2021)
Leukemia	D	Blood	Panel	32	963	0.84	0.88	0.94	Xu et al. (2016)
	Pg	Plasma/blood/BM	1 (*miR-155*)	11	1718	–	–	–	Zhang et al. (2018)
NSCLC	D	Plasma/serum/blood/tissue	1 (*miR-210*)	40	4741	0.72	0.8	0.83	He et al. (2018)
		Plasma/serum/blood		23	2321	0.65	0.76	0.77	
		Plasma		6	402	0.66	0.72	0.73	
		Serum		6	563	0.72	0.84	0.82	
		Blood		5	477	0.5	0.67	0.61	
		Sputum		6	879	0.68	0.79	0.81	
	Pg	Serum/BM/tissue	2	26	2653	–	–	–	Pop-Bica et al. (2018)
OC	D	Plasma/serum	23	33	1599	0.89	0.64	–	Wang et al. (2019)
			miR-200 family			0.9	0.58		
			1 (*miR-200c*)			0.85	0.52		
OC (EOC)	D	Serum/tissue	1 (*miR-200a*)	5	198	0.84	0.83	0.89	Teng et al. (2016)
			1 (*miR-200c*)	5		0.75	0.66	0.77	
OSCC	D	Plasma/serum/blood/saliva	22	16	2562	0.78	0.82	0.91	Rapado-González et al. (2019)
RCC	D	Plasma/serum	1 (*miR-378*)	12	1981	0.75	0.74	0.81	Li et al. (2016)

(continued)

Table 2.3 (continued)

Cancer	Biomark.	Sample origin	N° miRNAs	Studies	Subjects	Sensitivity	Specificity	AUC	References
Tyroid cancer	D	Plasma/serum/tissue	2 (*miR-221/222*)	24	2819	0.79	0.84	0.84	Liang et al. (2018)
Various cancers	D	Serum	1 (*miR-125b*)	8	1065	0.82	0.77	0.84	Wei et al. (2016)
		Plasma	2 (*miR-221/222*)	23	1147	0.76	0.75	0.82	Zhang et al. (2016)
		Plasma/serum	1 (*miR-31*)	14	2161	0.79	0.79	0.79	Ma et al. (2017)
		Plasma/serum/PBMCs	1 (*miR-21*)	48	5816	0.7	0.9	0.831	Gao et al. (2016a)
	Pg	Plasma/serum/tissue	1 (*miR-145*)	46	6875	–	–	–	Xu et al. (2019)
		Serum/tissue	2	27	3300	–	–	–	Peng and Duan (2019)

BC, Breast cancer; BCa, Bladder cancer; Biomark., Biomarker; BM, Bone marrow; CRC, colorectal cancer; CSF, cerebrospinal fluid; D, Diagnostic; DLBCL, diffuse large B-cell lymphoma; EOC, epithelial ovarian cancer; GC, gastric cancer; HCC, Hepatocellular carcinoma; HNSCC, Head and neck squamous cell carcinoma; OSCC, Oral squamous cell carcinoma; Pg, Prognostic; NSCLC, non-small cell lung cancer; OC, ovarian cancer; RCC, renal cell carcinoma

et al. 2017). In contrast, *miR-106* has been found to have a higher diagnostic accuracy in the plasma of gastric cancer patients than in serum (Peng et al. 2018). Other researchers found that circulating *miR-21* expression levels have high diagnostic accuracy in 14 different cancers analyzed (Gao et al. 2016a). Other studies analyzed the diagnostic/prognostic value of *miR-21* in mixed origin samples, bodily fluids, and tissues, and also found that is a promising biomarker in breast cancer, pancreatic cancer, and glioma, among others (Gao et al. 2016b; Jinling et al. 2017; Qu et al. 2016b; Zhao et al. 2020b). Other circulating miRNAs have been evaluated for their diagnostic capacity including a study of *miR-221/222* expression in plasma of 11 different cancer types, leading to a specificity of 0.76 and sensitivity of 0.75 with an AUC of 0.82 (Zhang et al. 2016). Furthermore, *miR-221/222* expression was also evaluated in tissue and bodily fluids in thyroid cancer patients, where it showed high accuracy in differentiating cancer patients from non-thyroid cancer patients (AUC: 0.88) (Liang et al. 2018). Other meta-analysis studies have analyzed the prognostic capacity of circulating miRNAs. For example levels of plasma *miR-130b* in HCC patients were found to correlate with disease-free survival (DFS) (Peng and Duan 2019) and with progression-free survival (PFS) in HCC, while levels of *miR-203* correlated with DFS and PFS in colorectal and breast cancer patients (Peng and Duan 2019; Shao et al. 2017). Similarly levels of *miR-155* and *miR-203* have been linked to overall survival in leukemia and CRC patients respectively (Shao et al. 2017; Zhang et al. 2018).

2.7 Conclusions

Since the discovery of circulating miRNAs as cancer biomarkers in 2008 (Lawrie et al. 2008), many studies have shown the potential usefulness of these molecules as noninvasive biomarkers. Of particular note, some miRNAs (and other ncRNAs) have been demonstrated to have better diagnostic capabilities than existing biomarkers (Izumi et al. 2021; Yang et al. 2021; Fredsøe et al. 2018). However, a lack of standardization including a great variety of methods for isolation, detection, and analysis of miRNAs precludes their uptake in routine clinical diagnostic practice (Larrea et al. 2016). In this regard, it is necessary to establish standard methodologies and reporting systems for circulating miRNA studies to make results robust between studies, in a similar manner that has been instigated with high-throughput methods such as microarray and NGS using the MIAME guidelines (Brazma et al. 2001). And beyond the reporting system standardization needs to be applied to the downstream computational analysis particularly when a range of predictive algorithms are used to impart functional significance without follow-on functional verification. Only when these issues are resolved can circulating miRNAs fulfil their clear potential and translate into routine clinical practice for the benefit of cancer patients.

References

Aad G, Abbott B, Abdallah J, Abdelalim AA, Abdesselam A, Abdinov O et al (2010) Search for new particles in two-jet final states in 7 TeV proton-proton collisions with the ATLAS detector at the LHC. Phys Rev Lett 105(16):161801

Afshar S, Afshar S, Warden E, Manochehri H, Saidijam M (2019) Application of artificial neural network in miRNA biomarker selection and precise diagnosis of colorectal cancer. Int J Mol Sci 23(3):175–183

Agarwal V, Bell GW, Nam JW, Bartel DP (2015) Predicting effective microRNA target sites in mammalian mRNAs. elife 4

Aguda BD, Kim Y, Piper-Hunter MG, Friedman A, Marsh CB (2008) MicroRNA regulation of a cancer network: consequences of the feedback loops involving miR-17-92, E2F, and Myc. Proc Natl Acad Sci U S A 105(50):19678–19683

Akers JC, Gonda D, Kim R, Carter BS, Chen CC (2013) Biogenesis of extracellular vesicles (EV): exosomes, microvesicles, retrovirus-like vesicles, and apoptotic bodies. J Neuro-Oncol 113(1): 1–11

Alix-Panabières C, Pantel K (2016) Clinical applications of circulating tumor cells and circulating tumor DNA as liquid biopsy. Cancer Discov 6(5):479–491

Arroyo JD, Chevillet JR, Kroh EM, Ruf IK, Pritchard CC, Gibson DF et al (2011) Argonaute2 complexes carry a population of circulating microRNAs independent of vesicles in human plasma. Proc Natl Acad Sci U S A 108(12):5003–5008

Atkin-Smith GK, Tixeira R, Paone S, Mathivanan S, Collins C, Liem M et al (2015) A novel mechanism of generating extracellular vesicles during apoptosis via a beads-on-a-string membrane structure. Nat Commun 6:7439

Aucher A, Rudnicka D, Davis DM (2013) MicroRNAs transfer from human macrophages to hepato-carcinoma cells and inhibit proliferation. J Immunol (Baltimore, Md: 1950) 191(12): 6250–6260

Ayupe AC, Tahira AC, Camargo L, Beckedorff FC, Verjovski-Almeida S, Reis EM (2015) Global analysis of biogenesis, stability and sub-cellular localization of lncRNAs mapping to intragenic regions of the human genome. RNA Biol 12(8):877–892

Bai X, Liu Z, Shao X, Wang D, Dong E, Wang Y et al (2019) The heterogeneity of plasma miRNA profiles in hepatocellular carcinoma patients and the exploration of diagnostic circulating miRNAs for hepatocellular carcinoma. PLoS One 14(2):e0211581

Bandyopadhyay S, Mitra R (2009) TargetMiner: microRNA target prediction with systematic identification of tissue-specific negative examples. Genome Biol 25(20):2625–2631

Baraniskin A, Nöpel-Dünnebacke S, Ahrens M, Jensen SG, Zöllner H, Maghnouj A et al (2013) Circulating U2 small nuclear RNA fragments as a novel diagnostic biomarker for pancreatic and colorectal adenocarcinoma. Int J Cancer 132(2):E48–E57

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116(2): 281–297

Bhattacharya A, Cui Y (2016) SomamiR 2.0: a database of cancer somatic mutations altering microRNA-ceRNA interactions. Nucleic Acids Res 44(D1):D1005–D1010

Bhattacharya A, Ziebarth JD, Cui Y (2013) SomamiR: a database for somatic mutations impacting microRNA function in cancer. Nucleic Acids Res 41(Database issue):D977–D982

Bhattacharya A, Ziebarth JD, Cui Y (2014) PolymiRTS database 3.0: linking polymorphisms in microRNAs and their target sites with human diseases and biological pathways. Nucleic Acids Res 42(Database issue):D86–D91

Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A (2018) Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 68(6):394–424

Brazma A, Hingamp P, Quackenbush J, Sherlock G, Spellman P, Stoeckert C et al (2001) Minimum information about a microarray experiment (MIAME)-toward standards for microarray data. Nat Genet 29(4):365–371

Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C et al (2019) The NHGRI-EBI GWAS catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. Nucleic Acids Res 47(D1):D1005–D1d12

Buschow SI, Nolte-'t Hoen EN, van Niel G, Pols MS, ten Broeke T, Lauwen M et al (2009) MHC II in dendritic cells is targeted to lysosomes or T cell-induced exosomes via distinct multivesicular body pathways. Traffic (Copenhagen, Denmark). 10(10):1528–1542

Bustos MA, Gross R, Rahimzadeh N (2020) A pilot study comparing the efficacy of lactate dehydrogenase levels versus circulating cell-free microRNAs in monitoring responses to checkpoint inhibitor immunotherapy in metastatic melanoma patients. Cancers (Basel) 12(11)

Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A et al (2011) Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. Genes Dev 25(18):1915–1927

Califf RM (2018) Biomarker definitions and their applications. Exp Biol Med (Maywood, NJ) 243(3):213–221

Carter JV, Galbraith NJ, Yang D, Burton JF, Walker SP, Galandiuk S (2017) Blood-based microRNAs as biomarkers for the diagnosis of colorectal cancer: a systematic review and meta-analysis. Sci Rep 116(6):762–774

Cavallari I, Grassi A, Del Bianco P, Aceti A, Zaborra C (2020) Prognostic stratification of bladder cancer patients with a microRNA-based approach. Cancers (Basel) 12(11)

Cerezo-Magaña M, Christianson HC, van Kuppevelt TH, Forsberg-Nilsson K, Belting M (2021) Hypoxic induction of exosome uptake through proteoglycan-dependent endocytosis fuels the lipid droplet phenotype in glioma. Mol Cancer Res 19(3):528–540

Chatrchyan S, Khachatryan V, Sirunyan AM, Tumasyan A, Adam W, Bergauer T et al (2011) Search for a vectorlike quark with charge 2/3 in t+Z events from pp collisions at sqrt[s]=7 TeV. Phys Rev Lett 107(27):271802

Chen X, Ba Y, Ma L, Cai X, Yin Y, Wang K et al (2008) Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. Cell Res 18(10): 997–1006

Chen X, Liang H, Zhang J, Zen K, Zhang CY (2012) Secreted microRNAs: a new form of intercellular communication. Trends Cell Biol 22(3):125–132

Chen S, Li T, Zhao Q, Xiao B, Guo J (2017) Using circular RNA hsa_circ_0000190 as a new biomarker in the diagnosis of gastric cancer. Clin Chim Acta 466:167–171

Cho S, Jang I, Jun Y, Yoon S, Ko M, Kwon Y et al (2013) MiRGator v3.0: a microRNA portal for deep sequencing, expression profiling and mRNA targeting. Nucleic Acids Res 41(Database issue):D252–D257

Christianson HC, Svensson KJ, van Kuppevelt TH, Li JP, Belting M (2013) Cancer cell exosomes depend on cell-surface heparan sulfate proteoglycans for their internalization and functional activity. Proc Natl Acad Sci U S A 110(43):17380–17385

Clancy JW, Zhang Y, Sheehan C, D'Souza-Schorey C (2019) An ARF6-Exportin-5 axis delivers pre-miRNA cargo to tumour microvesicles. Nat Cell Biol 21(7):856–866

Clark MB, Johnston RL, Inostroza-Ponta M, Fox AH, Fortini E, Moscato P et al (2012) Genome-wide analysis of long noncoding RNA stability. Genome Res 22(5):885–898

Cocucci E, Racchetti G, Meldolesi J (2009) Shedding microvesicles: artefacts no more. Trends Cell Biol 19(2):43–51

Cole H, Tran LT, Tran KD, Takeshima L, Stern SL, O'Day S et al (2019) Distinct methylation levels of mature microRNAs in gastrointestinal cancers. Cancers 10(1):3888

Colombo M, Raposo G, Théry C (2014) Biogenesis, secretion, and intercellular interactions of exosomes and other extracellular vesicles. Annu Rev Cell Dev Biol 30:255–289

Cullen BR (2004) Transcription and processing of human microRNA precursors. Mol Cell 16(6): 861–865

Dennis G Jr, Sherman BT, Hosack DA, Yang J, Gao W, Lane HC et al (2003) DAVID: database for annotation, visualization, and integrated discovery. Genome Biol 4(5):P3

Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A et al (2012) Landscape of transcription in human cells. Nature 489(7414):101–108

Dong L, Luo M, Wang F, Zhang J, Li T, Yu J (2013) TUMIR: an experimentally supported database of microRNA deregulation in various cancers. J Clin Bioinformatics 3(1):7

Duttagupta R, Jiang R, Gollub J, Getts RC, Jones KW (2011) Impact of cellular miRNAs on circulating miRNA biomarker signatures. PLoS One 6(6):e20769

Elias KM, Fendler W (2017) Diagnostic potential for a serum miRNA neural network for detection of ovarian cancer. Elife 6:e28932

Escola JM, Kleijmeer MJ, Stoorvogel W, Griffith JM, Yoshie O, Geuze HJ (1998) Selective enrichment of tetraspan proteins on the internal vesicles of multivesicular endosomes and on exosomes secreted by human B-lymphocytes. J Biol Chem 273(32):20121–20127

Escrevente C, Keller S, Altevogt P, Costa J (2011) Interaction and uptake of exosomes by ovarian cancer cells. BMC Cancer 11:108

Fabbri M, Paone A, Calore F, Galli R, Gaudio E, Santhanam R et al (2012) MicroRNAs bind to toll-like receptors to induce prometastatic inflammatory response. Proc Natl Acad Sci U S A 109(31):E2110–E2116

Fadok VA, Bratton DL, Frasch SC, Warner ML, Henson PM (1998) The role of phosphatidylserine in recognition of apoptotic cells by phagocytes. Cell Death Differ 5(7):551–562

Fang T, Lv H, Lv G, Li T, Wang C, Han Q et al (2018) Tumor-derived exosomal miR-1247-3p induces cancer-associated fibroblast activation to foster lung metastasis of liver cancer. Nat Commun 9(1):191

Fleischhacker M, Schmidt B (2010) Free circulating nucleic acids in plasma and serum (CNAPS) – useful for the detection of lung cancer patients? Cancer Biomarkers 6(3–4):211–219

Forberg D, Schwob T, Kempe R (2018) Catalytic condensation for the formation of polycyclic heteroaromatic compounds. Nat Commun 9(1):1751

Frank F, Sonenberg N, Nagar B (2010) Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2. Nature 465(7299):818–822

Fredsøe J, Rasmussen AKI, Thomsen AR, Mouritzen P, Høyer S, Borre M et al (2018) Diagnostic and prognostic microRNA biomarkers for prostate cancer in cell-free urine. Eur Urol Focus 4(6):825–833

Friard O, Re A, Taverna D, De Bortoli M, Corá D (2010) CircuitsDB: a database of mixed microRNA/transcription factor feed-forward regulatory circuits in human and mouse. BMC Bioinformatics 11:435

Fukada M, Matsuhashi N, Takahashi T, Sugito N, Heishima K, Yoshida K et al (2021) Postoperative changes in plasma miR21-5p as a novel biomarker for colorectal cancer recurrence: a prospective study. Cancer Sci 112(10):4270–4280

Gállego Pérez-Larraya J, Qu A, Wang W, Yang Y, Zhang X, Dong Y et al (2019) A serum piRNA signature as promising non-invasive diagnostic and prognostic biomarkers for colorectal cancer. Neuro-oncol Adv 11:3703–3720

Gallois-Montbrun S, Kramer B, Swanson CM, Byers H, Lynham S, Ward M et al (2007) Antiviral protein APOBEC3G localizes to ribonucleoprotein complexes found in P bodies and stress granules. J Virol 81(5):2165–2178

Gao Y, Dai M, Liu H, He W, Lin S, Yuan T et al (2016a) Diagnostic value of circulating miR-21: an update meta-analysis in various cancers and validation in endometrial cancer. Oncotarget 7(42): 68894–68908

Gao Y, Cai Q, Huang Y, Li S, Yang H, Sun L et al (2016b) MicroRNA-21 as a potential diagnostic biomarker for breast cancer patients: a pooled analysis of individual studies. Nat Commun 7(23):34498–34506

Geisler S, Coller J (2013) RNA in unexpected places: long non-coding RNA functions in diverse cellular contexts. Nat Rev Mol Cell Biol 14(11):699–712

Gibbings DJ, Ciaudo C, Erhardt M, Voinnet O (2009) Multivesicular bodies associate with components of miRNA effector complexes and modulate miRNA activity. Nat Cell Biol 11(9):1143–1149

Goodier JL, Zhang L, Vetter MR, Kazazian HH Jr (2007) LINE-1 ORF1 protein localizes in stress granules with other RNA-binding proteins, including components of RNA interference RNA-induced silencing complex. Mol Cell Biol 27(18):6469–6483

Gourzones C, Ferrand FR, Amiel C, Verillaud B, Barat A, Guerin M et al (2013) Consistent high concentration of the viral microRNA BART17 in plasma samples from nasopharyngeal carcinoma patients--evidence of non-exosomal transport. Virol J 10:119

Group F-NBW (2016) BEST (Biomarkers, EndpointS, and other Tools) Resource. Silver Spring (MD) Bethesda (MD): Food and Drug Administration (US) National Institutes of Health (US)

Gu X, Ma S, Liang B, Ju S (2021) Serum hsa_tsr016141 as a kind of tRNA-derived fragments is a novel biomarker in gastric cancer. Front Oncol 11:679366

Guo Z, Maki M, Ding R, Yang Y, Zhang B, Xiong L (2014) Genome-wide survey of tissue-specific microRNA and transcription factor regulatory networks in 12 tissues. Sci Rep 4:5150

Guo H, Chitiprolu M, Roncevic L, Javalet C, Hemming FJ, Trung MT et al (2017) Atg5 disassociates the V(1)V(0)-ATPase to promote exosome production and tumor metastasis independent of canonical macroautophagy. Nat Cell Biol 43(6):716–30.e7

Guo X, Lv X, Ru Y, Zhou F, Wang N, Xi H et al (2020) Circulating exosomal gastric cancer-associated Long noncoding RNA1 as a biomarker for early detection and monitoring progression of gastric cancer: a multiphase study. JAMA Surg 155(7):572–579

Gurung S, Perocheau D, Touramanidou L, Baruteau J (2021) The exosome journey: from biogenesis to uptake and intracellular signalling. Cell Commun Signal 19(1):47

Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D et al (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. Nature 458(7235):223–227

Hallal S, Ebrahim Khani S, Wei H, Lee MYT, Sim HW, Sy J et al (2020) Deep sequencing of small RNAs from neurosurgical extracellular vesicles substantiates miR-486-3p as a circulating biomarker that distinguishes glioblastoma from lower-grade astrocytoma patients. Int J Mol Sci 21(14)

Han J, Sun W, Liu R, Zhou Z, Zhang H, Chen X et al (2020) Plasma exosomal miRNA expression profile as oxaliplatin-based chemoresistant biomarkers in colorectal adenocarcinoma. Front Oncol 10:1495

Harris MA, Clark J, Ireland A, Lomax J, Ashburner M, Foulger R et al (2004) The gene ontology (GO) database and informatics resource. Nucleic Acids Res 32(Database issue):D258–D261

He RQ, Cen WL, Cen JM, Cen WN, Li JY, Li MW et al (2018) Clinical significance of miR-210 and its prospective signaling pathways in non-small cell lung cancer: evidence from gene expression omnibus and the cancer genome Atlas data mining with 2763 samples and validation via real-time quantitative PCR. Cell Physiol Biochem 46(3):925–952

Heijnen HF, Schiel AE, Fijnheer R, Geuze HJ, Sixma JJ (1999) Activated platelets release two types of membrane vesicles: microvesicles by surface shedding and exosomes derived from exocytosis of multivesicular bodies and alpha-granules. Blood 94(11):3791–3799

Heneghan HM, Miller N, Kerin MJ (2010) Circulating miRNA signatures: promising prognostic tools for cancer. J Clin Oncol 28(29):e573–4; author reply e5–6

Hobor F, Dallmann A, Ball NJ (2018) A cryptic RNA-binding domain mediates Syncrip recognition and exosomal partitioning of miRNA targets. Nat Commun 9(1):831

Hoshino A, Costa-Silva B, Shen TL, Rodrigues G, Hashimoto A, Tesic Mark M et al (2015) Tumour exosome integrins determine organotropic metastasis. Nature 527(7578):329–335

Hsu PW, Huang HD, Hsu SD, Lin LZ, Tsou AP, Tseng CP et al (2006) miRNAMap: genomic maps of microRNA genes and their target genes in mammalian genomes. Nucleic Acids Res 34 (Database issue):D135–D139

Hsu JB, Chiu CM, Hsu SD, Huang WY, Chien CH, Lee TY et al (2011) miRTar: an integrated system for identifying miRNA-target interactions in human. BMC Bioinformatics. 12:300

Hu J, Wang Z, Tan CJ, Liao BY, Zhang X, Xu M et al (2013) Plasma microRNA, a potential biomarker for acute rejection after liver transplantation. Transplantation 95(8):991–999

Hua X, Tang R, Xu X, Wang Z, Xu Q, Chen L et al (2018) mirTrans: a resource of transcriptional regulation on microRNAs for human cell lines. Nucleic Acids Res 46(D1):D168–Dd74

Huang JT, Liu SM, Ma H, Yang Y, Zhang X, Sun H et al (2016) Systematic review and meta-analysis: circulating miRNAs for diagnosis of hepatocellular carcinoma. Cancer Sci 231(2): 328–335

Huang HY, Lin YC, Li J, Huang KY, Shrestha S, Hong HC et al (2020) miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. Nucleic Acids Res 48 (D1):D148–Dd54

Iavello A, Frech VS, Gai C, Deregibus MC, Quesenberry PJ, Camussi G (2016) Role of Alix in miRNA packaging during extracellular vesicle biogenesis. Int J Mol Med 37(4):958–966

Imani S, Zhang X, Hosseinifard H, Fu S, Fu J (2017) The diagnostic role of microRNA-34a in breast cancer: a systematic review and meta-analysis. Oncotarget 8(14):23177–23187

Imaoka H, Toiyama Y, Fujikawa H, Hiro J, Saigusa S, Tanaka K et al (2016) Circulating microRNA-1290 as a novel diagnostic and prognostic biomarker in human colorectal cancer. Ann Oncol 27(10):1879–1886

Ismail N, Wang Y, Dakhlallah D, Moldovan L, Agarwal K, Batte K et al (2013) Macrophage microvesicles induce macrophage differentiation and miR-223 transfer. Blood 121(6):984–995

Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y et al (2015) The landscape of long noncoding RNAs in the human transcriptome. Nat Genet 47(3):199–208

Izumi D, Zhu Z, Chen Y, Toden S, Huo X, Kanda M et al (2021) Assessment of the diagnostic efficiency of a liquid biopsy assay for early detection of gastric cancer. JAMA Netw Open 4(8): e2121129

Janas T, Janas T, Yarus M (2006) Specific RNA binding to ordered phospholipid bilayers. Nucleic Acids Res 34(7):2128–2136

Janas T, Janas T, Yarus M (2012) Human tRNA(Sec) associates with HeLa membranes, cell lipid liposomes, and synthetic lipid bilayers. Int J Mol Sci 18(12):2260–2268

Janas T, Janas MM, Sapoń K, Janas T (2015) Mechanisms of RNA loading into exosomes. FEBS Lett 589(13):1391–1398

Jeggari A, Marks DS, Larsson E (2012) miRcode: a map of putative microRNA target sites in the long non-coding transcriptome. Bioinformatics (Oxford, England). 28(15):2062–2063

Jeppesen DK, Fenix AM, Franklin JL, Higginbotham JN, Zhang Q, Zimmerman LJ et al (2019) Reassessment of exosome composition. Cell 177(2):428–45.e18

Jiang Q, Wang Y, Hao Y, Juan L, Teng M, Zhang X et al (2009) miR2Disease: a manually curated database for microRNA deregulation in human disease. Nucleic Acids Res 37(Database issue): D98–104

Jinling W, Sijing S, Jie Z, Guinian W (2017) Prognostic value of circulating microRNA-21 for breast cancer: a systematic review and meta-analysis. Artif Cells Nanomed Biotechnol 45(6): 1–6

John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS (2004) Human MicroRNA targets. PLoS Biol 2(11):e363

Johnstone RM, Adam M, Hammond JR, Orr L, Turbide C (1987) Vesicle formation during reticulocyte maturation. Association of plasma membrane activities with released vesicles (exosomes). J Biol Chem 262(19):9412–9420

Jordens I, Fernandez-Borja M, Marsman M, Dusseljee S, Janssen L, Calafat J et al (2001) The Rab7 effector protein RILP controls lysosomal transport by inducing the recruitment of dynein-dynactin motors. Curr Biol 11(21):1680–1685

Kang JW, Eun YG (2021) Diagnostic value of salivary miRNA in head and neck squamous cell cancer: systematic review and meta-analysis. 22(13)

Kang M, Ren M, Li Y, Fu Y, Deng M, Li C (2018) Exosome-mediated transfer of lncRNA PART1 induces gefitinib resistance in esophageal squamous cell carcinoma via functioning as a competing endogenous RNA. J Exp Clin Cancer Res 37(1):171

Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I et al (2018) DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. Nucleic Acids Res 46(D1):D239–Dd45

Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E (2007) The role of site accessibility in microRNA target recognition. Nat Genet 39(10):1278–1284

Khorshid M, Hausser J, Zavolan M, van Nimwegen E (2013) A biophysical miRNA-mRNA interaction model infers canonical and noncanonical targets. Nat Methods 10(3):253–255

Kirschner MB, Kao SC, Edelman JJ, Armstrong NJ, Vallely MP, van Zandwijk N et al (2011) Haemolysis during sample preparation alters microRNA content of plasma. PLoS One 6(9): e24145

Kitagawa T, Taniuchi K, Tsuboi M, Sakaguchi M, Kohsaki T, Okabayashi T et al (2019) Circulating pancreatic cancer exosomal RNAs for detection of pancreatic cancer. J Cell Mol Med 13(2):212–227

Kleene KC (2018) Y-box proteins combine versatile cold shock domains and arginine-rich motifs (ARMs) for pleiotropic functions in RNA biology. Biochem J 475(17):2769–2784

Köhler J, Schuler M, Gauler TC, Nöpel-Dünnebacke S, Ahrens M, Hoffmann AC et al (2016) Circulating U2 small nuclear RNA fragments as a diagnostic and prognostic biomarker in lung cancer patients. J Cancer Res Clin Oncol 142(4):795–805

Konishi H, Kubota T, Shiozaki A, Fujiwara H, Okamoto K, Otsuji E et al (2020) Exosomal transfer of miR-15b-3p enhances tumorigenesis and malignant transformation through the DYNLT1/Caspase-3/Caspase-9 signaling pathway in gastric cancer. Sci Rep 39(1):32

Kosaka N, Iguchi H, Ochiya T (2010a) Circulating microRNA in body fluid: a new potential biomarker for cancer diagnosis and prognosis. Cancer Sci 101(10):2087–2092

Kosaka N, Iguchi H, Yoshioka Y, Takeshita F, Matsuki Y, Ochiya T (2010b) Secretory mechanisms and intercellular transfer of microRNAs in living cells. J Biol Chem 285(23): 17442–17452

Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ et al (2005) Combinatorial microRNA target predictions. Nat Genet 37(5):495–500

Kuhlmann JD, Baraniskin A, Hahn SA, Mosel F, Bredemeier M, Wimberger P et al (2014) Circulating U2 small nuclear RNA fragments as a novel diagnostic tool for patients with epithelial ovarian cancer. Clin Chem 60(1):206–213

Lai NS, Wu DG, Fang XG, Lin YC, Chen SS, Li ZB et al (2015) Serum microRNA-210 as a potential noninvasive biomarker for the diagnosis and prognosis of glioma. Br J Cancer 112(7): 1241–1246

Lambowitz AM, Kirino Y, Mourelatos Z (2007) The mouse homolog of HEN1 is a potential methylase for Piwi-interacting RNAs. Proc Natl Acad Sci U S A 13(9):1397–1401

Larrea E, Sole C, Manterola L, Goicoechea I, Armesto M, Arestin M et al (2016) New concepts in cancer biomarkers: circulating miRNAs in liquid biopsies. Int J Mol Sci 17(5)

Laulagnier K, Javalet C, Hemming FJ, Chivet M, Lachenal G, Blot B et al (2018) Amyloid precursor protein products concentrate in a subset of exosomes specifically endocytosed by neurons. Cellular Mol Life Sci 75(4):757–773

Lawrie CH, Gal S, Dunlop HM, Pushkaran B, Liggins AP, Pulford K et al (2008) Detection of elevated levels of tumour-associated microRNAs in serum of patients with diffuse large B-cell lymphoma. Br J Haematol 141(5):672–675

Leidal AM, Huang HH (2020) The LC3-conjugation machinery specifies the loading of RNA-binding proteins into extracellular vesicles. Nat Cell Biol 22(2):187–199

Li JH, Liu S, Zhou H, Qu LH, Yang JH (2014a) starBase v2.0: decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res 42(Database issue):D92–D97

Li Y, Qiu C, Tu J, Geng B, Yang J, Jiang T et al (2014b) HMDD v2.0: a database for experimentally supported human microRNA and disease associations. Nucleic Acids Res 42(Database issue): D1070–D1074

Li ZZ, Shen LF, Li YY, Chen P, Chen LZ (2016) Clinical utility of microRNA-378 as early diagnostic biomarker of human cancers: a meta-analysis of diagnostic test. Oncotarget 7(36): 58569–58578

Li Z, Niu H, Qin Q, Yang S, Wang Q, Yu C et al (2019a) lncRNA UCA1 mediates resistance to cisplatin by regulating the miR-143/FOSL2-signaling pathway in ovarian cancer. Mol Therapy Nucleic Acids 17:92–101

Li Z, Zhou Y, Yang G, He S, Qiu X, Zhang L et al (2019b) Using circular RNA SMARCA5 as a potential novel biomarker for hepatocellular carcinoma. Clin Chim Acta 492:37–44

Li JR, Tong CY, Sung TJ, Kang TY, Zhou XJ, Liu CC (2019c) CMEP: a database for circulating microRNA expression profiling. Bioinformatics (Oxford, England). 35(17):3127–3132

Liang L, Zheng X, Hu M, Cui Y, Zhong Q, Wang S et al (2018) MiRNA-221/222 in thyroid cancer: a meta-analysis. Clin Chim Acta 484:284–292

Liang ZX, Liu HS, Wang FW, Xiong L, Zhou C, Hu T et al (2019) LncRNA RPPH1 promotes colorectal cancer metastasis by interacting with TUBB3 and by promoting exosomes-mediated macrophage M2 polarization. Cell Death Dis 10(11):829

Liang H, Jiao Z, Rong W, Qu S, Liao Z, Sun X et al (2020) 3'-terminal 2'-O-methylation of lung cancer miR-21-5p enhances its stability and association with Argonaute 2. Nucleic Acids Res 48(13):7027–7040

Liao J, Yu L, Mei Y, Guarnera M, Shen J, Li R et al (2010) Small nucleolar RNA signatures as biomarkers for non-small-cell lung cancer. Cancer Cell Int 9:198

Lin CC (2021) Extracellular vesicle miR-200c enhances gefitinib sensitivity in heterogeneous EGFR-mutant NSCLC. J Immunotherapy Cancer 9(3)

Lin LY, Yang L, Zeng Q, Wang L, Chen ML, Zhao ZH et al (2018) Tumor-originated exosomal lncUEGC1 as a circulating biomarker for early-stage gastric cancer. Mol Cancer 17(1):84

Liu W, Wang X (2019) Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. Genome Biol 20(1):18

Liu H, Yue D, Chen Y, Gao SJ, Huang Y (2010) Improving performance of mammalian microRNA target prediction. BMC Bioinformatics 11:476

Liu T, Zhang X, Gao S, Jing F, Yang Y, Du L et al (2016) Exosomal long noncoding RNA CRNDE-h as a novel serum-based biomarker for diagnosis and prognosis of colorectal cancer. Oncotarget 7(51):85551–85563

Lo KW, Lo YM, Leung SF, Tsang YS, Chan LY, Johnson PJ et al (1999) Analysis of cell-free Epstein-Barr virus associated RNA in the plasma of patients with nasopharyngeal carcinoma. Clin Chem 45(8 Pt 1):1292–1294

López JI, Pulido R, Cortés JM, Angulo JC, Lawrie CH (2018) Potential impact of PD-L1 (SP-142) immunohistochemical heterogeneity in clear cell renal cell carcinoma immunotherapy. Pathol Res Pract 214(8):1110–1114

Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D et al (2005) MicroRNA expression profiles classify human cancers. Nature 435(7043):834–838

Ma Y, Chen Y, Lin J, Liu Y, Luo K, Cao Y et al (2017) Circulating miR-31 as an effective biomarker for detection and prognosis of human cancer: a meta-analysis. Oncotarget 8(17): 28660–28671

Mailman MD, Feolo M, Jin Y, Kimura M, Tryka K, Bagoutdinov R et al (2007) The NCBI dbGaP database of genotypes and phenotypes. Nat Genet 39(10):1181–1186

Mandel P, Metais P (1948) Nuclear acids in human blood plasma. Comptes rendus des seances de la Societe de biologie et de ses filiales 142(3–4):241–243

Mańka R, Janas P, Sapoń K, Janas T, Janas T (2021) Role of RNA motifs in RNA interaction with membrane lipid rafts: implications for therapeutic applications of exosomal RNAs. Int J Mol Sci 22(17)

Marks LS, Fradet Y, Deras IL, Blase A, Mathis J, Aubin SM et al (2007) PCA3 molecular urine assay for prostate cancer in men undergoing repeat biopsy. Urology 69(3):532–535

McDonald JS, Milosevic D, Reddi HV, Grebe SK, Algeciras-Schimnich A (2011) Analysis of circulating microRNA: preanalytical and analytical challenges. Clin Chem 57(6):833–840

Melé M, Mattioli K, Mallard W, Shechner DM, Gerhardinger C, Rinn JL (2017) Chromatin environment, transcriptional regulation, and splicing distinguish lincRNAs and mRNAs. Genome Res 27(1):27–37

Melo SA, Sugimoto H, O'Connell JT, Kato N, Villanueva A, Vidal A et al (2014) Cancer exosomes perform cell-independent microRNA biogenesis and promote tumorigenesis. Nat Cell Biol 26(5):707–721

Michela B (2021) Liquid biopsy: a family of possible diagnostic tools. Diagnostics (Basel, Switzerland) 11(8)

Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, Pogosova-Agadjanyan EL et al (2008) Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U S A 105(30):10513–10518

Mitra R, Bandyopadhyay S (2011) MultiMiTar: a novel multi objective optimization based miRNA-target prediction method. PLoS One 6(9):e24583

Mittelbrunn M, Gutiérrez-Vázquez C, Villarroya-Beltri C, González S, Sánchez-Cabo F, González M et al (2011) Unidirectional transfer of microRNA-loaded exosomes from T cells to antigen-presenting cells. Nat Cell Biol 2:282

Mjelle R, Dima SO, Bacalbasa N, Chawla K, Sorop A, Cucu D et al (2019) Comprehensive transcriptomic analyses of tissue, serum, and serum exosomes from hepatocellular carcinoma patients. BMC Cancer 19(1):1007

Muralidharan-Chari V, Clancy J, Plou C, Romao M, Chavrier P, Raposo G et al (2009) ARF6-regulated shedding of tumor cell-derived plasma membrane microvesicles. Curr Biol 19(22): 1875–1885

Murillo OD, Thistlethwaite W, Rozowsky J, Subramanian SL, Lucero R, Shah N et al (2019) exRNA Atlas analysis reveals distinct extracellular RNA cargo types and their carriers present across human biofluids. Cell 177(2):463–77.e15

Naeem H, Küffner R, Csaba G, Zimmer R (2010) miRSel: automated extraction of associations between microRNAs and genes from the biomedical literature. BMC Bioinformatics 11:135

Nishibeppu K, Komatsu S, Imamura T (2020) Plasma microRNA profiles: identification of miR-1229-3p as a novel chemoresistant and prognostic biomarker in gastric cancer. Oncogene 10(1):3161

Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M (1999) KEGG: Kyoto encyclopedia of genes and genomes. Nucleic Acids Res 27(1):29–34

Ohara T, Sakaguchi Y, Suzuki T, Ueda H, Miyauchi K, Suzuki T (2007) The 3′ termini of mouse Piwi-interacting RNAs are 2'-O-methylated. Nat Struct Mol Biol 14(4):349–350

Panagiotara A, Markou A, Lianidou ES, Patrinos GP, Katsila T (2017) Exosomes: a cancer theranostics road map. Public Health Genomics 20(2):116–125

Paraskevopoulou MD, Georgakilas G, Kostoulas N, Vlachos IS, Vergoulis T, Reczko M et al (2013) DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. Nucleic Acids Res 41(Web Server issue):W169–W173

Peng Z, Duan F (2019) Prognostic values of microRNA-130 family expression in patients with cancer: a meta-analysis and database test. J Transl Med 17(1):347

Peng Q, Zhang X, Min M, Zou L, Shen P, Zhu Y (2017) The clinical role of microRNA-21 as a promising biomarker in the diagnosis and prognosis of colorectal cancer: a systematic review and meta-analysis. Oncotarget 8(27):44893–44909

Peng Q, Shen Y, Lin K, Zou L, Shen Y, Zhu Y (2018) Comprehensive and integrative analysis identifies microRNA-106 as a novel non-invasive biomarker for detection of gastric cancer. J Transl Med 16(1):127

Peng XX, Yu R, Wu X, Wu SY, Pi C, Chen ZH et al (2020) Correlation of plasma exosomal microRNAs with the efficacy of immunotherapy in EGFR / ALK wild-type advanced non-small cell lung cancer. J Immunother Cancer 8(1)

Perdikopanis N, Georgakilas GK, Grigoriadis D, Pierros V, Kavakiotis I, Alexiou P et al (2021) DIANA-miRGen v4: indexing promoters and regulators for more than 1500 microRNAs. Nucleic Acids Res 49(D1):D151–D1d9

Pezuk JA, Miller TLA, Bevilacqua JLB, de Barros A, de Andrade FEM, LFA EM et al (2017) Measuring plasma levels of three microRNAs can improve the accuracy for identification of malignant breast lesions in women with BI-RADS 4 mammography. elife 8(48):83940–83948

Pigati L, Yaddanapudi SC, Iyengar R, Kim DJ, Hearn SA, Danforth D et al (2010) Selective release of microRNA species from normal and malignant mammary epithelial cells. PLoS One 5(10): e13515

Ponjavic J, Ponting CP, Lunter G (2007) Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. Genome Res 17(5):556–565

Poon IK, Chiu YH, Armstrong AJ, Kinchen JM, Juncadella IJ, Bayliss DA et al (2014) Unexpected link between an antibiotic, pannexin channels and apoptosis. Nature 507(7492):329–334

Pop-Bica C, Pintea S, Cojocneanu-Petric R, Del Sal G, Piazza S, Wu ZH et al (2018) MiR-181 family-specific behavior in different cancers: a meta-analysis view. Cancer Metastasis Rev 37(1):17–32

Pritchard CC, Kroh E, Wood B, Arroyo JD, Dougherty KJ, Miyaji MM et al (2012) Blood cell origin of circulating microRNAs: a cautionary note for cancer biomarker studies. Cancer Prev Res 5(3):492–497

Progida C, Spinosa MR, De Luca A, Bucci C (2006) RILP interacts with the VPS22 component of the ESCRT-II complex. Biochem Biophys Res Commun 347(4):1074–1079

Puigdelloses M, González-Huárriz M, García-Moure M, Martínez-Vélez N, Esparragosa Vázquez I, Bruna J et al (2020) RNU6–1 in circulating exosomes differentiates GBM from non-neoplastic brain lesions and PCNSL but not from brain metastases. Neuro-oncol Adv 2(1):vdaa010

Qu L, Ding J, Chen C, Wu ZJ, Liu B, Gao Y et al (2016a) Exosome-transmitted lncARSR promotes Sunitinib resistance in renal cancer by acting as a competing endogenous RNA. Cancer Cell 29(5):653–668

Qu K, Lin T, Pang Q, Liu T, Wang Z, Tai M et al (2016b) Extracellular miRNA-21 as a novel biomarker in glioma: evidence from meta-analysis, clinical validation and experimental investigations. Oncotarget 7(23):33994–34010

Qu A, Wang W, Yang Y, Zhang X, Dong Y, Zheng G et al (2019) A serum piRNA signature as promising non-invasive diagnostic and prognostic biomarkers for colorectal cancer. Cancer Manag Res 11:3703–3720

Rapado-González O, Martínez-Reglero C, Salgado-Barreira A, López-López R, Suárez-Cunqueiro MM, Muinelo-Romay L (2019) miRNAs in liquid biopsy for oral squamous cell carcinoma diagnosis: systematic review and meta-analysis. Oral Oncol 99:104465

Raposo G, Nijman HW, Stoorvogel W, Liejendekker R, Harding CV, Melief CJ et al (1996) B lymphocytes secrete antigen-presenting vesicles. J Exp Med 183(3):1161–1172

Rennie W, Liu C, Carmack CS, Wolenc A, Kanoria S, Lu J et al (2014) STarMir: a web server for prediction of microRNA binding sites. Nucleic Acids Res 42(Web Server issue):W114–W118

Rennie W, Kanoria S, Liu C, Mallick B, Long D, Wolenc A et al (2016) STarMirDB: a database of microRNA binding sites. RNA Biol 13(6):554–560

Richter VA, Semenov DV, Baraniskin A, Nöpel-Dünnebacke S, Schumacher B, Gerges C et al (2014) Analysis of U2 small nuclear RNA fragments in the bile differentiates cholangiocarcinoma from primary sclerosing cholangitis and other benign biliary disorders. Biomed Res Int 59(7):1436–1441

Rupaimoole R, Slack FJ (2017) MicroRNA therapeutics: towards a new era for the management of cancer and other diseases. Nat Rev Drug Discov 16(3):203–222

Russo F, Di Bella S, Nigita G, Macca V, Laganà A, Giugno R et al (2012) miRandola: extracellular circulating microRNAs database. PLoS One 7(10):e47786

Santangelo L, Giurato G, Cicchini C, Montaldo C, Mancone C, Tarallo R et al (2016) The RNA-binding protein SYNCRIP is a component of the hepatocyte exosomal machinery controlling MicroRNA sorting. Cell Rep 17(3):799–808

Savill J, Hogg N, Ren Y, Haslett C (1992) Thrombospondin cooperates with CD36 and the vitronectin receptor in macrophage recognition of neutrophils undergoing apoptosis. J Clin Invest 90(4):1513–1522

Schaefer CF, Anthony K, Krupa S, Buchoff J, Day M, Hannay T et al (2009) PID: the pathway interaction database. Nucleic Acids Res 37(Database issue):D674–D679

Schully SD, Yu W, McCallum V, Benedicto CB, Dong LM, Wulf A et al (2011) Cancer GAMAdb: database of cancer genetic associations from meta-analyses and genome-wide association studies. Eur J Hum Genetics 19(8):928–930

Schwarzenbach H, Nishida N, Calin GA, Pantel K (2014) Clinical relevance of circulating cell-free microRNAs in cancer. Nat Rev Clin Oncol 11(3):145–156

Shao Y, Gu W, Ning Z, Song X, Pei H, Jiang J (2017) Evaluating the prognostic value of microRNA-203 in solid tumors based on a meta-analysis and the cancer genome Atlas (TCGA) datasets. Cell Physiol Biochem 41(4):1468–1480

Sharova E, Bertazzolo I, D'Agostino DM, Iafrate M, Ciminale V, Kim MY et al (2021) Urinary exosomal microRNA profiling in intermediate-risk prostate cancer. Cancers. 11(1):7355

Shukla V, Varghese VK, Kabekkodu SP, Mallya S, Chakrabarty S, Jayaram P et al (2019) Enumeration of deregulated miRNAs in liquid and tissue biopsies of cervical cancer. PLoS One 155(1):135–143

Shukuya T, Ghai V, Amann JM, Okimoto T, Shilo K, Kim TK et al (2020) Circulating MicroRNAs and extracellular vesicle-containing microRNAs as response biomarkers of anti-programmed cell death protein 1 or programmed death-ligand 1 therapy in NSCLC. J Thoracic Oncol 15(11): 1773–1781

Shurtleff MJ, Yao J, Qin Y, Nottingham RM, Temoche-Diaz MM, Schekman R (2017) Broad role for YBX1 in defining the small noncoding RNA composition of exosomes. Proc Natl Acad Sci USA 114(43):E8987–E8e95

Smith MA, Gesell T, Stadler PF, Mattick JS (2013) Widespread purifying selection on RNA structure in mammals. Nucleic Acids Res 41(17):8220–8236

Solé C, Lawrie CH (2021) MicroRNAs in metastasis and the tumour microenvironment. Int J Mol Sci 22(9):4859

Sole C, Arnaiz E, Manterola L, Otaegui D, Lawrie CH (2019) The circulating transcriptome as a source of cancer liquid biopsy biomarkers. Semin Cancer Biol 58:100–108

Sorenson GD, Pribish DM, Valone FH, Memoli VA, Bzik DJ, Yao SL (1994) Soluble normal and mutated DNA sequences from single-copy genes in human blood. Cancer Epidemiol Biomarkers Prevent 3(1):67–71

Sork H, Corso G, Krjutskov K, Johansson HJ (2018) Heterogeneity and interplay of the extracellular vesicle small RNA transcriptome and proteome. Sci Rep 8(1):10813

Su J, Liao J, Gao L, Shen J, Guarnera MA, Zhan M et al (2016) Analysis of small nucleolar RNAs in sputum for lung cancer diagnosis. Oncotarget 7(5):5131–5142

Sun J, Sun Z, Gareev I, Yan T, Chen X, Ahmad A et al (2021) Exosomal miR-2276-5p in plasma is a potential diagnostic and prognostic biomarker in glioma. Front Cell Dev Biol 9:671202

Sur D, Balacescu L, Cainap SS, Visan S, Pop L, Burz C et al (2021) Predictive efficacy of MiR-125b-5p, MiR-17-5p, and MiR-185-5p in liver metastasis and chemotherapy response among advanced stage colorectal cancer patients. Front Oncol 11:651380

Takizawa F, Tsuji S, Nagasawa S (1996) Enhancement of macrophage phagocytosis upon iC3b deposition on apoptotic cells. FEBS Lett 397(2–3):269–272

Tan C, Cao J, Chen L, Xi X, Wang S, Zhu Y et al (2019) Noncoding RNAs serve as diagnosis and prognosis biomarkers for hepatocellular carcinoma. Clin Chem 65(7):905–915

Tauro BJ, Greening DW, Mathias RA, Ji H, Mathivanan S, Scott AM et al (2012) Comparison of ultracentrifugation, density gradient separation, and immunoaffinity capture methods for isolating human colon cancer cell line LIM1863-derived exosomes. Methods (San Diego, Calif) 56(2):293–304

Tavares RCA, Pyle AM, Somarowthu S (2019) Phylogenetic analysis with improved parameters reveals conservation in lncRNA structures. J Mol Biol 431(8):1592–1603

Temoche-Diaz MM, Shurtleff MJ (2019) Distinct mechanisms of microRNA sorting into cancer cell-derived extracellular vesicle subtypes. Elife 8

Teng Y, Su X, Zhang X, Zhang Y, Li C, Niu W et al (2016) miRNA-200a/c as potential biomarker in epithelial ovarian cancer (EOC): evidence based on miRNA meta-signature and clinical investigations. Oncotarget 7(49):81621–81633

Tian T, Zhu YL, Zhou YY, Liang GF, Wang YY, Hu FH et al (2014) Exosome uptake through clathrin-mediated endocytosis and macropinocytosis and mediating miR-21 delivery. J Biol Chem 289(32):22258–22267

Ting CY, Liew SM, Price A, Gan GG, Bee-Lan Ong D, Tan SY et al (2019) Clinical significance of aberrant microRNAs expression in predicting disease relapse/refractoriness to treatment in diffuse large B-cell lymphoma: a meta-analysis. Crit Rev Oncol Hematol 144:102818

Toda Y, Takata K, Nakagawa Y, Kawakami H, Fujioka S, Kobayashi K et al (2015) Effective internalization of U251-MG-secreted exosomes into cancer cells and characterization of their lipid components. Biochem Biophys Res Commun 456(3):768–773

Tong YS, Wang XW, Zhou XL, Liu ZH, Yang TX, Shi WH et al (2015) Identification of the long non-coding RNA POU3F3 in plasma as a novel biomarker for diagnosis of esophageal squamous cell carcinoma. Mol Cancer 14:3

Tong Z, Cui Q, Wang J, Zhou Y (2019) TransmiR v2.0: an updated transcription factor-microRNA regulation database. Nucleic Acids Res 47(D1):D253–D2d8

Trajkovic K, Hsu C, Chiantia S, Rajendran L, Wenzel D, Wieland F et al (2008) Ceramide triggers budding of exosome vesicles into multivesicular endosomes. Science (New York, NY). 319(5867):1244–1247

Tsai CC, Chen TY, Tsai KJ, Lin MW, Hsu CY, Wu DC et al (2020) NF-κB/miR-18a-3p and miR-4286/BZRAP1 axis may mediate carcinogenesis in *Helicobacter pylori*-associated gastric cancer. Biomed Pharmacother 132:110869

Turchinovich A, Weiz L, Langheinz A, Burwinkel B (2011) Characterization of extracellular circulating microRNA. Nucleic Acids Res 39(16):7223–7233

Turchinovich A, Weiz L, Burwinkel B (2012) Extracellular miRNAs: the mystery of their origin and function. Trends Biochem Sci 37(11):460–465

Umu SU, Langseth H (2018) A comprehensive profile of circulating RNAs in human serum. RNA Biol 15(2):242–250

Valadi H, Ekstrom K, Bossios A, Sjostrand M, Lee JJ, Lotvall JO (2007a) Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. Nat Cell Biol 9(6):654–659

Valadi H, Ekström K, Bossios A, Sjöstrand M, Lee JJ, Lötvall JO (2007b) Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. Cell Rep 9(6):654–659

van den Berg MMJ, Krauskopf J, Ramaekers JG, Kleinjans JCS, Prickaerts J, Briedé JJ (2020) Circulating microRNAs as potential biomarkers for psychiatric and neurodegenerative disorders. Nat Commun 185:101732

van Engeland M, Nieland LJ, Ramaekers FC, Schutte B, Reutelingsperger CP (1998) Annexin V-affinity assay: a review on an apoptosis detection system based on phosphatidylserine exposure. Cytometry 31(1):1–9

Vasioukhin V, Anker P, Maurice P, Lyautey J, Lederrey C, Stroun M (1994) Point mutations of the N-ras gene in the blood plasma DNA of patients with myelodysplastic syndrome or acute myelogenous leukaemia. Br J Haematol 86(4):774–779

Vastrik I, D'Eustachio P, Schmidt E, Gopinath G, Croft D, de Bono B et al (2007) Reactome: a knowledge base of biologic pathways and processes. Genome Biol 8(3):R39

Vickers KC, Palmisano BT, Shoucri BM, Shamburek RD, Remaley AT (2011) MicroRNAs are transported in plasma and delivered to recipient cells by high-density lipoproteins. Nat Cell Biol 13(4):423–433

Villarroya-Beltri C, Gutiérrez-Vázquez C, Sánchez-Cabo F, Pérez-Hernández D, Vázquez J, Martin-Cofreces N et al (2013) Sumoylated hnRNPA2B1 controls the sorting of miRNAs into exosomes through binding to specific motifs. Nat Commun 4:2980

Vychytilova-Faltejskova P, Stitkovcova K, Radova L, Sachlova M, Kosarova Z, Slaba K et al (2018) Circulating PIWI-interacting RNAs piR-5937 and piR-28876 are promising diagnostic biomarkers of colon cancer. Cancer Epidemiol Biomarkers Preven 27(9):1019–1028

Wagner J, Riwanto M, Besler C, Knau A, Fichtlscherer S, Röxe T et al (2013) Characterization of levels and cellular transfer of circulating lipoprotein-bound microRNAs. Arterioscler Thromb Vasc Biol 33(6):1392–1400

Wang T, Hong W (2006) RILP interacts with VPS22 and VPS36 of ESCRT-II and regulates their membrane recruitment. Biochem Biophys Res Commun 350(2):413–423

Wang K, Zhang S, Weber J, Baxter D, Galas DJ (2010) Export of microRNAs and microRNA-protective protein by mammalian cells. Nucleic Acids Res 38(20):7248–7259

Wang D, Gu J, Wang T, Ding Z (2014) OncomiRDB: a database for the experimentally verified oncogenic and tumor-suppressive microRNAs. Bioinformatics (Oxford, England). 30(15): 2237–2238

Wang S, Li W, Lian B, Liu X, Zhang Y, Dai E et al (2015) TMREC: a database of transcription factor and MiRNA regulatory cascades in human diseases. PLoS One 10(5):e0125222

Wang YH, Ji J, Weng H, Wang BC, Wang FB (2018) MiR-139 in digestive system tumor diagnosis and detection: bioinformatics and meta-analysis. Clin Chim Acta 485:33–41

Wang X, Kong D, Wang C, Ding X, Zhang L, Zhao M et al (2019) Circulating microRNAs as novel potential diagnostic biomarkers for ovarian cancer: a systematic review and updated meta-analysis. J Ovarian Res 12(1):24

Wang D, Wang X, Si M, Yang J, Sun S, Wu H et al (2020a) Exosome-encapsulated miRNAs contribute to CXCL12/CXCR4-induced liver metastasis of colorectal cancer by enhancing M2 polarization of macrophages. Cancer Sci 474:36–52

Wang J, Ma G, Li M, Han X, Xu J, Liang M et al (2020b) Plasma tRNA fragments derived from 5′ ends as novel diagnostic biomarkers for early-stage breast cancer. Mol Therapy Nucleic Acids 21:954–964

Weber JA, Baxter DH, Zhang S, Huang DY, Huang KH, Lee MJ et al (2010) The microRNA spectrum in 12 body fluids. Clin Chem 56(11):1733–1741

Wei X, Chen D, Lv T, Li G, Qu S (2016) Serum microRNA-125b as a potential biomarker for glioma diagnosis. Mol Neurobiol 53(1):163–170

Wood MJA, Mäger I, El Andaloussi S, Yao B, La LB, Chen YC et al (2012) Defining a new role of GW182 in maintaining miRNA stability. Sci Rep 13(12):1102–1108

Wozniak AL, Adams A, King KE, Dunn W, Christenson LK, Hung WT et al (2020) The RNA binding protein FMR1 controls selective exosomal miRNA cargo loading during inflammation. Cancer Cell 219(10)

Wu SY, Huang YT, Chang WY, Su PL, Su WC, Zhang WC et al (2016) Tumour-initiating cell-specific miR-1246 and miR-1290 expression converge to promote non-small cell lung cancer progression. Biomedicine 7:11702

Wu B, Su S, Patil DP, Liu H, Gan J, Jaffrey SR (2018) Molecular basis for the specific and multivariant recognitions of RNA substrates by human hnRNP A2/B1. elife 9(1):420

Wubbolts R, Leckie RS, Veenhuizen PT, Schwarzmann G, Möbius W, Hoernschemeyer J et al (2003) Proteomic and biochemical analyses of human B cell-derived exosomes. Potential implications for their function and multivesicular body formation. J Biol Chem 278(13): 10963–10972

Wulfken LM, Moritz R, Ohlmann C, Holdenrieder S, Jung V, Becker F et al (2011) MicroRNAs in renal cell carcinoma: diagnostic implications of serum miR-1233 levels. PLoS One 6(9):e25787

Xia X, Wang S, Ni B, Xing S, Cao H, Zhang Z et al (2020) Hypoxic gastric cancer-derived exosomes promote progression and metastasis via MiR-301a-3p/PHD3/HIF-1α positive feedback loop. Oncogene 39(39):6231–6244

Xiao Y, Zhang L, Song Z, Guo C, Zhu J, Li Z et al (2016) Potential diagnostic and prognostic value of plasma circulating microRNA-182 in human glioma. Med Sci Monit 22:855–862

Xie B, Ding Q, Han H, Wu D (2013) miRCancer: a microRNA-cancer association database constructed by text mining on literature. Bioinformatics (Oxford, England). 29(5):638–644

Xie L, Yao Z, Zhang Y, Li D, Hu F, Liao Y et al (2018) Deep RNA sequencing reveals the dynamic regulation of miRNA, lncRNAs, and mRNAs in osteosarcoma tumorigenesis and pulmonary metastasis. Cell Death Dis 9(7):772

Xu J, Chen Q, Zen K, Zhang C, Zhang Q (2013) Synaptosomes secrete and uptake functionally active microRNAs via exocytosis and endocytosis pathways. Int J Mol Sci 124(1):15–25

Xu LH, Guo Y, Zhang XL, Chen JJ, Hu SY (2016) Blood-based circulating MicroRNAs are potential diagnostic biomarkers for leukemia: result from a meta-analysis. Int J Mol Sci 38(3): 939–949

Xu L, Zhang Y, Tang J, Wang P, Li L, Yan X et al (2019) The prognostic value and regulatory mechanisms of microRNA-145 in various tumors: a systematic review and meta-analysis of 50 studies. Cancer Epidemiol Biomark Prev 28(5):867–881

Yang J, Li X, Wei S, Peng L, Sang H, Jin D et al (2021) Evaluation of the diagnostic potential of a plasma exosomal miRNAs panel for gastric cancer. Front Oncol 11:683465

Yau TO, Tang CM, Harriss EK, Dickins B (2019) Faecal microRNAs as a non-invasive tool in the diagnosis of colonic adenomas and colorectal cancer: a meta-analysis 9(1):9491

Yin J, Feng Y, Yang Z, Shang J, Qu K, Lin T et al (2016) Extracellular miRNA-21 as a novel biomarker in glioma: evidence from meta-analysis, clinical validation and experimental investigations. J Transl Med 7(23):33994–34010

Yu T, Wang Y, Fan Y, Fang N, Wang T, Xu T et al (2019) CircRNAs in cancer metabolism: a review. J Hematol Oncol 12(1):90

Zhang R, Pang B, Xin T, Guo H, Xing Y, Xu S et al (2016) Plasma miR-221/222 family as novel descriptive and prognostic biomarkers for glioma. Mol Neurobiol 53(3):1452–1460

Zhang X, Wang Y, Guo Q, Diao Y, Liu H, Song G et al (2018) Prognostic role of microRNA-155 in patients with leukemia: a meta-analysis. Clin Chim Acta 483:6–13

Zhang Z, Yin J, Lu C, Wei Y, Zeng A, You Y (2019) Exosomal transfer of long non-coding RNA SBF2-AS1 enhances chemoresistance to temozolomide in glioblastoma. J Exp Clin Cancer Res 38(1):166

Zhao R, Zhang Y, Zhang X, Yang Y, Zheng X, Li X et al (2018a) Exosomal long noncoding RNA HOTTIP as potential novel diagnostic and prognostic biomarker test for gastric cancer. Mol Cancer 17(1):68

Zhao SY, Wang J, Ouyang SB, Huang ZK, Liao L (2018b) Salivary circular RNAs Hsa_Circ_0001874 and Hsa_Circ_0001971 as novel biomarkers for the diagnosis of Oral squamous cell carcinoma. Cell Physiol Biochem 47(6):2511–2521

Zhao Y, Yan Y, Ma R, Lv X, Zhang L, Wang J et al (2020a) Expression signature of six-snoRNA serves as novel non-invasive biomarker for diagnosis and prognosis prediction of renal clear cell carcinoma. J Cell Mol Med 24(3):2215–2228

Zhao F, Wei C, Cui MY, Xia QQ, Wang SB, Zhang Y (2020b) Prognostic value of microRNAs in pancreatic cancer: a meta-analysis. Aging 12(10):9380–9404

Zheng Q, Chen C, Guan H, Kang W, Yu C (2017) Prognostic role of microRNAs in human gastrointestinal cancer: a systematic review and meta-analysis. Oncotarget 8(28):46611–46623

Zheng M, Hou L, Ma Y, Zhou L, Wang F, Cheng B et al (2019) Exosomal let-7d-3p and miR-30d-5p as diagnostic biomarkers for non-invasive screening of cervical cancer and its precursors. Mol Cancer 18(1):76

Zhou W, Fong MY, Min Y, Somlo G, Liu L, Palomares MR et al (2014) Cancer-secreted miR-105 destroys vascular endothelial barriers to promote metastasis. Cancer Cell 25(4):501–515

Zhou KR, Liu S, Sun WJ, Zheng LL, Zhou H, Yang JH et al (2017) ChIPBase v2.0: decoding transcriptional regulatory networks of non-coding RNAs and protein-coding genes from ChIP-seq data. Nucleic Acids Res 45(D1):D43–d50

Zhou Q, Liu J, Quan J, Liu W, Tan H, Li W (2018) MicroRNAs as potential biomarkers for the diagnosis of glioma: a systematic review and meta-analysis. Cancer Sci 109(9):2651–2659

Zhou L, Wang W, Wang F, Yang S, Hu J, Lu B et al (2021) Plasma-derived exosomal miR-15a-5p as a promising diagnostic biomarker for early detection of endometrial carcinoma. Mol Cancer 20(1):57

Zhu L, Li T, Shen Y, Yu X, Xiao B, Guo J (2019) Using tRNA halves as novel biomarkers for the diagnosis of gastric cancer. Cancer Biomarkers 25(2):169–176

Zietzer A, Hosen MR, Wang H, Goody PR, Sylvester M, Latz E et al (2020) The RNA-binding protein hnRNPU regulates the sorting of microRNA-30c-5p into large extracellular vesicles. J Extracellular Vesicles 9(1):1786967

Chapter 3
Regulation of Immune Cells by microRNAs and microRNA-Based Cancer Immunotherapy

Rafaela Nasser Veiga, Érika Pereira Zambalde, Livia Cox, Tayana Shultz Jucoski, Ana Flávia Kohler, Tamyres Mingorance Carvalho, Ana Carolina Rodrigues, Beatriz Ludwig, Kaitlyn Crowley, Jaqueline Carvalho de Oliveira, and Maria Angelica Cortez

Abstract MicroRNAs (miRNAs) are small (~21 nucleotides) endogenous noncoding RNA molecules involved in the posttranscriptional regulation of gene expression. Modulation of gene expression by miRNAs occurs via base-pairing of the specific miRNA primary sequence to its corresponding target messenger RNA, which can be located either in the 3′ untranslated region or within the coding sequence. This pairing can lead to either translational repression or cleavage of the mRNA, resulting in reduced levels of the target protein. MiRNAs are involved in mediating and controlling several interactions between immune and cancer cells and are also important regulators of immune responses. Increasing interest has focused on elucidating the role of miRNAs in the regulation of anticancer immune responses and how this could affect the efficacy of different cancer therapeutics. Indeed, immune responses have both pro- and anti-oncogenic effects, and functional interactions between immune and cancer cells in the tumor microenvironment are crucial in determining the course of cancer progression. Thus, understanding the role of miRNAs in controlling cancer immunity is important for revealing mechanisms that could be modulated to enhance the success of immunotherapy for patients with

R. N. Veiga · T. S. Jucoski · A. F. Kohler · T. M. Carvalho · A. C. Rodrigues · B. Ludwig · J. C. de Oliveira
Laboratory of Human Cytogenetics and Oncogenetics, Postgraduate Program in Genetics, Department of Genetics, Universidade Federal do Paraná, Rua Coronel Francisco Heráclito dos Santos, Curitiba, Brazil

É. P. Zambalde
Multidisciplinary Laboratory of Food and Health, School of Applied Sciences, State University of Campinas, Limeira, SP, Brazil

L. Cox · K. Crowley · M. A. Cortez (✉)
Department of Radiation Oncology, The University of Texas MD Anderson Cancer Center, Houston, TX, USA
e-mail: macortez@mdanderson.org

U. Schmitz et al. (eds.), *Systems Biology of MicroRNAs in Cancer*, Advances in Experimental Medicine and Biology 1385,
https://doi.org/10.1007/978-3-031-08356-3_3

cancer. In this chapter, we discuss the involvement of miRNAs in the regulation of immune cells and potential therapeutic approaches in which miRNAs are used for cancer immunotherapy.

Keywords miRNA · Cancer · Tumor microenvironment · Immunotherapy · Immune cells · Immune checkpoint molecules

3.1 Introduction

Over the past few decades, substantial efforts have been made by the scientific community to reveal the mechanisms of cancer immune surveillance and determine how tumor cells acquire the ability to evade the immune system (Drake et al. 2006; Swann and Smyth 2007; Chow et al. 2012; Spranger and Gajewski 2018). Understanding these mechanisms could facilitate the discovery of new druggable targets and therapeutic approaches to improve response to anticancer treatments.

The goal of the growing field of cancer immunotherapy is to boost immune system function by activating and strengthening individual patients' own immune systems to efficiently fight tumors. New immunotherapeutic approaches are being developed, and some are already being applied in the clinic. One of the best-known immunotherapy strategies to date involves the use of monoclonal antibodies that target cancer cells by recognizing and blocking immune checkpoint molecules expressed on their surfaces such as cytotoxic T-lymphocyte-associated protein 4 (CTLA4) and programmed cell death 1 (PD1) (Wei et al. 2018). Other strategies involve the use of chimeric antigen receptor (CAR) T-cell therapy (Neelapu et al. 2017), cancer vaccines (Wang et al. 2018c), immunostimulatory cytokines (Rosenberg et al. 2015), and oncolytic viruses, which can be classified as both immune and biological therapy (Pol et al. 2016).

Unfortunately, only a subset of cancer patients treated with immunotherapy have experienced positive clinical outcomes, in part because of immune-related adverse events in diverse biological systems that can lead to severe toxicity during treatment (Rashdan et al. 2018; Dahiya et al. 2020; De Martin et al. 2020; Harrison et al. 2020; Presotto et al. 2020; Muntyanu et al. 2021). Therefore, identifying those patients who would benefit from immunotherapy without experiencing major immune-related adverse events is essential for avoiding unnecessary toxicity, particularly among those who would not efficiently respond to immunotherapeutic strategies. Characterizing the tumor microenvironment (TME), including its immune cell infiltrates and soluble signaling components, is of utmost importance for identifying potential biomarkers of response (or lack of response) to immunotherapy.

MicroRNAs (miRNAs) are short, single-stranded RNA molecules that are usually 20–23 nucleotides long (Croce and Calin 2005; Farazi et al. 2011). These noncoding RNA molecules participate in the regulation of gene expression through posttranscriptional repression of mRNAs, which are targeted by strong complementarity between mRNA sequence and miRNA seed sequences (Bartel 2009).

The expression of miRNAs is often dysregulated during cancer development and progression (Sathipati and Ho 2018; Angius et al. 2019; Fehlmann et al. 2020), which consequently dysregulates the expression of many coding genes that control

important biological processes in cancer (Angius et al. 2019). These findings make miRNAs interesting potential targets for cancer therapy. Because miRNAs are involved in regulating immune responses during cancer progression (Okada et al. 2010; Hirschberger et al. 2018; Omar et al. 2019), miRNA-based therapies might also enrich the field of immunotherapy.

Therefore, the scientific community has been seeking to identify the mechanisms by which miRNAs regulate the immune responses against cancer and how dysregulation of miRNAs could influence various cancer therapies. Although one function of the immune system is to provide constant surveillance for identifying and eliminating cancer cells, the immune system can also have pro-oncogenic effects. For example, inflammatory responses may spur the production of factors that enhance tumor cell growth and survival as well as angiogenesis in addition to prompting the secretion of soluble molecules that promote the activation of epithelial–mesenchymal transition pathways (Hanahan and Weinberg 2011). Thus, understanding how miRNAs could affect the complex interactions between immune cells and cancer cells in the TME is crucial to developing effective, specific, and safe miRNA-based therapeutic strategies against cancer.

In this chapter, we review the role of miRNAs in the regulation of immune cells and discuss the potential for, and challenges with, miRNA-based therapies against cancer.

3.2 Innate Immunity

The innate immune response can be crucial in tumor prevention, initiation, and progression. The outcome of pro- versus antitumor innate immune responses depends on the TME, which first primes and then reinforces the differentiation and response of infiltrating cells, cytokines, chemokines, matrix-degrading enzymes, and growth factors present in situ (Curtale 2018). The most abundant infiltrating cells are macrophages, myeloid-derived suppressor cells (MDSCs), neutrophils, mast cells, dendritic cells (DCs), and natural killer (NK) cells (Xing et al. 2021).

Tumor regulatory networks also involve miRNAs, which can coordinate different immune response functions in the TME or be transferred from one cell to another via exosomes (Lone et al. 2021). These miRNAs may also regulate the expression of immunomodulatory molecules in both tumor and immune cells in ways that facilitate tumor immune escape (Hirschberger et al. 2018). In this section, we focus on the role of miRNAs in the innate immune response, specifically their immunomodulatory effects on macrophages, DCs, and NK cells. These various roles are discussed in the following paragraphs and are summarized in Table 3.1.

3.2.1 Macrophages

Macrophages are a heterogeneous group of cells with a wide spectrum of activation states, ranging from the classically activated (M1 macrophages) to the alternatively

Table 3.1 Role of miRNAs in tumor immune response

miRNA	Targets	Immune response	References
let-7a/ let-7i	SOCS1	Suppresses immune evasion; induces macrophage M2 polarization; reduces maturation and functional state of DCs; promotes tumor escape	Zhang et al. (2011), Park et al. (2019)
miR-130a	TGF-βR	Promotes antitumor immunity; restricts tumor metastasis	Bader (2012)
miR-138	CTLA4; PD1; PDL1	Regulates immune checkpoints	Wei et al. (2016), Rasoolnezhad et al. (2021)
miR-142	IRF8; CD40L; TRAF6; IRAK1; PDL1	Regulates maturation process, pro-inflammatory cytokine secretion, and TCD4 differentiation of DCs; decreases NF-kB signaling; enhances antitumor immunity	Belz (2013), Fordham et al. (2015), Naqvi et al. (2015), Jia et al. (2017), Berrien-Elliott et al. (2019), Taghikhani et al. (2019)
miR-145	TGF-βR	Promotes antitumor immunity; restricts tumor metastasis	Bader (2012)
miR-146a/ miR-146b	TRAF6; IRAK1; STAT1; STAT3; MyD88	Modulates TLR4 signaling; decreases NF-kB signaling; modulates Th1 response	Taganov et al. (2006), Nahid et al. (2009), Chen et al. (2011), El Gazzar et al. (2011), Du et al. (2012), Curtale et al. (2013), Zhang et al. (2013), Park et al. (2015), Xu et al. (2017a), Simanovich et al. (2018), Wang et al. (2018b)
miR-149/ miR-149-3p	CSF-1; PD1; TIM-3; BTLA; FOXP1	Blocks paracrine interactions with macrophages; reverses CD8+ T-cell exhaustion	Zhang et al. (2019), Sánchez-González et al. (2020)
miR-150/ miR-150-5p	Prf1; PIK3AP1; AKT2	Facilitates cytotoxic capacity and maturation of NK cells; reduces maturation of DCs; decreases IL-10	Bezman et al. (2010), Kim et al. (2011, 2014), Tung et al. (2018)
miR-152	–	Reduces maturation of DCs, reduces IL-6 and IL-12	Smyth et al. (2015), Tung et al. (2018)
miR-155	c-Fos; TLR; IL13RA; LPS/INF-y; SHIP-1; TIM-3; c-Maf; SOCS1; PDL1	Reduces maturation and functional state of DCs; regulates NF-κB signaling; induces macrophage M1 polarization; suppresses INF-y production; contributes to thymic Treg cell differentiation; regulates Th1 and Th17 response; sensitizes	Eis et al. (2005), Rodriguez et al. (2007), Kohlhaas et al. (2009), Dunand-Sauthier et al. (2011), Martinez-Nunez et al. (2011), Cai et al. (2012), Chang et al. (2012), Ji et al. (2015), Mashima (2015), Huffaker et al. (2017), Yee

(continued)

Table 3.1 (continued)

miRNA	Targets	Immune response	References
		B-lymphoma cells to anti-PDL1 antibody	et al. (2017), Wang et al. (2018a), Chen et al. (2020), Monnot et al. (2020), Dong et al. (2021)
miR-15a	FOXP3; CTLA4	Regulates T cell cycle and memory T cell differentiation	Liu et al. (2014)
miR-17-92 cluster	TGF-βRII; CREB1; PTEN; RORα; PHLPP2	Inhibits cell surface molecules on DCs; enhances DC endocytosis; regulates Th1 response; enhances IFN-γ production; suppresses and regulates T cell differentiation; accelerates P13K signaling	Jiang et al. (2011), de Kouchkovsky et al. (2013), Ohno et al. (2013), Ranji et al. (2013), Kosaka et al. (2015), Kuo et al. (2019)
miR-181	NLK; Smad7; PIAS3	Promotes NK differentiation; reduces IFN-γ translation; increases TGF-β-induced signaling; inhibits Th17 cell differentiation	Zhang et al. (2018c), Jiang et al. (2020)
miR-183	DAP12	Inhibits NK cytotoxicity	Donatelli et al. (2014)
miR-186	TGF-β1	Inhibits NK cytotoxicity	Neviani et al. (2019)
miR-20	MICA/B	Inhibits NK cytotoxicity	Zhu et al. (2018)
miR-200a/ miR-200b/ miR-200c	CD47 (B6H12); PDL1	Promotes phagocytosis of macrophages; regulates immune checkpoints	Chen et al. (2014), Rigoutsos et al. (2017), Katakura et al. (2020)
miR-203	MHC II; RUNX1	Reduces antigen-presenting and regulatory capabilities of DCs; induces macrophage M2 polarization	Zhou et al. (2014), Shinohara et al. (2017), Takano et al. (2017)
miR-21	TLR; DUSP10; PIAS3; STAT3	Activates NF-κB signaling; induces macrophage M2 polarization; reduces IL-12; regulates CD8+ cell proliferation and the polarization and inflammatory responses of Th1 and Th2 through the IL-2 and IFN-γ pathways; contributes to immune cell recruitment	Kwak et al. (2011), Chang et al. (2012), Fabbri et al. (2012), Wang et al. (2013b, 2015b), Okoye et al. (2014), Smyth et al. (2015), Tung et al. (2018), Hsieh et al. (2018), Anastasiadou et al. (2019), Ren et al. (2019a), Syed et al. (2019), Hong et al. (2020)
miR-212-3p	MHC II; RFXAP	Reduces antigen-presenting and regulatory capabilities of DCs; inhibits immune tolerance	Ding et al. (2015)
miR-214	JAK/STAT; PTEN	Suppresses macrophage M2 polarization	Okoye et al. (2014), Gao et al. (2020)

(continued)

Table 3.1 (continued)

miRNA	Targets	Immune response	References
miR-24	STAT5; FGF11	Promotes Treg cells and differentiation into Th1 and Th17 cells	Fordham et al. (2015), Naqvi et al. (2015), Ma et al. (2016), Ye et al. (2016), Zhang et al. (2018a)
miR-27a-3p/ miR-27a-5p	ICOS; Prf1; GzmB	Promotes antitumor immunity; reduces cytotoxic capacity and maturation of NK cells	Kim et al. (2011), Ma et al. (2016), Yao et al. (2020)
miR-29	TLR	Activates NF-κB signaling, which increases the expression of pro-inflammatory cytokines	Fabbri et al. (2012), Nygren et al. (2014), Ren et al. (2019a)
miR-34a/ miR-34a-5p	PD1/PDL1	Increases immunogenicity in mixed lymphocyte reactions; regulates immune checkpoints	Bader (2012), Yang et al. (2012), Cortez et al. (2016), Anastasiadou et al. (2019), Hong et al. (2020)
miR-374b	PD1	Regulates immune checkpoints	Huang et al. (2018)
miR-410-5p	miR-410-3p	Forms a duplex with -3p arm and promotes degradation, thereby suppressing its ability to inhibit tumor angiogenesis	Wang et al. (2017)
miR-424	PDL1	Regulates immune checkpoints; activates T cell response	Xu et al. (2016)
miR-5119	PDL1; IDOS	Regulates immune checkpoints	Zhang et al. (2020)

Abbreviations: BMDMs, bone marrow-derived macrophages; CD40L, cluster of differentiation 40 ligand; CREB1, CAMP-responsive element binding protein 1; CSF-1, colony-stimulating factor-1; CTLA4, cytotoxic T-lymphocyte associated protein 4; DAP12, DNAX-activating protein 12 kDa; DCs, dendritic cells; DUSP10, dual-specificity phosphatase 10; FGF11, fibroblast growth factor 11; FOXP1, forkhead box P1; FOXP3, forkhead box P3; GzmB, granzyme B; IL-10, interleukin-10; IL-12, interleukin-12; IL-13RA, interleukin-13 receptor alfa; IL-6, interleukin-6; INF-γ, interferon-gamma; IRAK1, interleukin-1 receptor-associated kinase 1; IRF8, interferon regulatory factor 8; JAK, Janus kinase; MHC, major histocompatibility complex; NF-kB, nuclear factor kappa beta; NK, natural-killer cells; NLK, Notch signaling inhibitor Nemo-like kinase; PIAS3, protein inhibitor of activated STAT3; PD1: programmed cell death protein 1; PDL1: programmed death-ligand 1; PI3K, phosphoinositide-3-kinase; PTEN: phosphatase and tensin homolog; Prf1, perforin 1; RFXAP, regulatory factor X-associated protein; RORα, RAR-related orphan receptor alpha; RUNX1, runt-related transcription factor 1; SHIP-1, SH2-containing inositol 5′-phosphatase 1; SOCS1, suppressor of cytokine signaling 1; STAT, signal transducer and activator of transcription; TAMs, tumor-associated macrophages; Tim-3, T-cell immunoglobulin, mucin domain-3; TGF-β1, transforming growth factor beta; TGF-BRII, transforming growth factor beta receptor 2; TLR, Toll-like receptors; TRAF6, TNF receptor-associated factor 6; Treg, regulatory T cell

activated (M2 macrophages) (Sica et al. 2015). In the early stages of tumor development, pro-inflammatory M1 macrophages predominate and result in antitumor activity (Qin et al. 2012). However, chronic inflammation alters the TME so as to promote the alternatively activated M2 macrophages, otherwise known as anti-inflammatory, pro-tumorigenic, or tumor-associated macrophages (TAMs) (Zhang et al. 2017; Syed et al. 2019). Indeed, TAMs are the major inflammation-related component of the TME (Gordon and Martinez 2010).

In addition to regulating macrophage polarization, miRNAs can also target several critical signaling pathways in macrophages associated with tumor-associated inflammation (Syed et al. 2019). MiR-21 and miR-29 bind to the toll-like receptors (TLRs) on TAMs and activate NF-κB signaling, which increases the expression of pro-inflammatory cytokines (Syed et al. 2019). Also, enhanced delivery of miR-21-5p by extracellular vesicles (EVs) promotes macrophage M2 polarization after hypoxia in lung cancer (Ren et al. 2019a) and epithelial–mesenchymal transition-mediated M2 polarization in Snail overexpressing cancer cells (Hsieh et al. 2018).

MiR-155 is also associated with M1 polarization: it silences the interleukin-13 receptor A (IL-13RA), stabilizes tumor necrosis factor (TNF)-α, and targets CEBP/β, thereby blocking M2 polarization (Martinez-Nunez et al. 2011). MiR-155 can also prompt M1 polarization in bone marrow-derived macrophages (BMDMs) upon exposure to lipopolysaccharide (LPS) or interferon (IFN)-γ (Cai et al. 2012).

On the other hand, M2 polarization in BMDMs occurs in response to miR-146a, which is induced by IL-4 (Zhang et al. 2013). The main targets of miR-146a are IL-1 receptor-associated kinase 1 (IRAK1) and TNF receptor-associated factor 6 (TRAF6) adaptor molecules in the TLR pathway (Taganov et al. 2006). MiR-146a participates in a negative feedback loop that dampens the signals upstream of NF-κB (Nahid et al. 2009). MiR-146a also promotes interactions between RBM4 and Ago2, which prompt the assembly of the miRNA-induced silencing complex, which disrupts translation of TNF-α in THP-1 monocytic cells (El Gazzar et al. 2011). On the other hand, the anti-inflammatory activity of IL-10 relies in part on the induction of miR-146b in a STAT3-dependent manner. MiR-146b also modulates TLR4 signaling by dampening MyD88, IRAK1, and TRAF6 signaling (Curtale et al. 2013).

Other miRNAs have distinct functions in macrophage polarization (Table 3.1). For example, transfer of the miRNA let-7a by EVs under hypoxic tumor conditions promotes M2 polarization (Park et al. 2019). miR-149 impairs the infiltration of primary tumors by M2 macrophages (Sánchez-González et al. 2020), and miR-214 suppresses M2 polarization via inhibiting JAK2/STAT3 signaling (Gao et al. 2020).

3.2.2 Dendritic Cells

DCs are antigen-presenting cells with the ability to regulate adaptive immunity. Mature DCs promote the activation, proliferation, and differentiation of effector T cells, whereas immature DCs downregulate T-cell responses (Morelli and Thomson 2007). DCs acquire tumor-specific or tumor-associated antigens and are activated by

the recognition of activator signals arising from stressed or dying cancer cells (Yatim et al. 2017). The capacity to induce a tumor-specific immune response in vivo correlates with the degree of DC maturation (Labeur et al. 1999).

A variety of miRNAs including miR-155, let-7i, miR-22, miR-21, miR-146, miR-34a, and miR-142 have been implicated in DC differentiation and maturation. Mechanistically, miR-142 upregulates IFN regulatory factor 8 (IRF8), which regulates the differentiation of DCs into $CD4^+$ DCs (Belz 2013). MiRNA-146a regulates the maturation of DCs and pro-inflammatory cytokine secretion by DCs by targeting CD40L, TRAF6, and IRAK1, which decrease NF-κB signaling (Chen et al. 2011; Park et al. 2015). Inhibition of let-7i depresses the maturation and functional state of DCs via targeting suppressor of cytokine signaling 1 (SOCS1) (Zhang et al. 2011). MiR-155 silences c-Fos expression, another crucial factor in DC maturation and function (Dunand-Sauthier et al. 2011). Overexpression of miR-17-5p inhibits the expression of cell surface molecules on DCs and enhances DC endocytosis (Cui et al. 2019).

DCs also use exosomal miRNAs as an efficient mechanism of intercellular communication (Montecalvo et al. 2012). Exosomal delivery of let-7i and miR-142 regulates the maturation of DCs and promotes tumor escape through a variety of mechanisms (Taghikhani et al. 2019). Intercellular communications also have substantial influence on Tregs, which transfer EVs containing miRNA to antigen-presenting cells during immune recognition, which leads to immune modulation (Mittelbrunn et al. 2011). In one study, high levels of miR-21, miR-148, and miR-152 were found to be present in EVs derived from Tregs. One mechanism by which miRNAs are transferred is by modifying the cytokine profile of DCs (Smyth et al. 2015). MiR-21 inhibits IL-12p35 production, whereas miR-148 and miR-152 suppress IL-6 and IL-12 production, both of which are cytokines involved in DC maturation (Tung et al. 2018). Decreases in IL-12, IL-6, and TNF-α after TLR blockade in DCs were also observed after ectopic expression of miR-142-3p (Fordham et al. 2015; Naqvi et al. 2015). Moreover, miR-150-5p may be linked to the increased IL-10 levels seen in DCs treated with Treg-derived EVs (Tung et al. 2018).

The complex communications between the TME and DCs refine the regulation of cancer cells via miRNAs. For example, pre-miR-410 can be processed into miR-410-5p in cancer cells and miR-410-3p in DCs. Cancer cells stimulate the DCs to produce miR-410-3p, which can inhibit angiogenesis by targeting vascular endothelial growth factor (VEGF)-α. However, cancer cells also express miR-410-5p, which can be transported into DCs where they form a duplex with miR-410-3p. The duplex formation promotes miR-410-3p degradation and thus suppresses its ability to inhibit tumor angiogenesis (Wang et al. 2017).

3.2.3 Natural Killer Cells

NK cells identify, target, and kill cancer cells. NK cells can also coordinate both the innate and the adaptive immune responses against foreign pathogens and

transformed cells (Yu et al. 2013). NK cells express various surface receptors that either promote or inhibit cell killing by those NKs. Generally, the predominance of activating vs. inhibiting receptor signaling is responsible for the induction of NK-mediated killing (Fabbri 2020).

In various animal and human cell culture systems, the miRNA profile can also affect NK cell development and function. Disruption of the global miRNA profile in mouse NK cells by the deletion of either Dicer or Dgcr8, both important to miRNAs biogenesis, resulted in decreased NK cell survival, maturation, and proliferation (Bezman et al. 2010; Sullivan et al. 2013). Other microarray studies revealed that highly expressed conserved miRNAs are present in both mouse and human NK cells (Bezman et al. 2010).

MiRNAs can also affect the cytotoxic capacity of NK cells via miR-150, which targets the perforin 3′ UTR, and miR-27a*-5p, which targets *Prf1* and *GzmB* expression (Kim et al. 2011, 2014). Other miRNAs such as miR-146a negatively regulate IFN-γ production in NK cells by targeting IRAK1 and TRAF6, with subsequent inhibition of the NF-κB signaling cascade; miR-146a also reduces the production of TNF-α by targeting STAT1, therefore reducing NK cell-mediated cytotoxicity (Xu et al. 2017a; Wang et al. 2018b).

Most of the endogenous miRNAs that have been characterized so far have been shown to modulate NK cell antitumor activity in the TME (Pesce et al. 2020). TGF-β, a key mediator in the TME, increases the expression of mature miRNA-1245, which acts to block NKG2D-mediated immune responses in NK cells and thus support the TME (Espinoza et al. 2012). TGF-β also induces miR-183, which abrogates the tumor cell killing function of NK cells by targeting DNAX-activating protein of 12 kDa (DAP12) (Donatelli et al. 2014).

The EVs released by NK cells also transport miRNAs that are capable of mounting strong antitumor effects in the TME (Fabbri 2020). In one study, lung cancer-derived EVs containing miRNAs suppressed the antitumor activities of NK cells by targeting CD107a expression (Berchem et al. 2016). The delivery of miR-186 to MYCN-amplified NK cells was found to impair their survival and migration and prevent the TGF-β1-dependent inhibition of NK cytotoxicity, thereby altering the cytotoxic potential of NKs (Neviani et al. 2019).

3.3 Adaptive [Acquired] Immunity

3.3.1 T-Helper Cells

MiRNAs control many aspects of the adaptive immune response through their regulation of T-helper cells. These key mediators of immune function become activated upon interacting with antigens expressed on the surface of antigen-presenting cells (e.g., macrophages or DCs), which prompts an intracellular signaling cascade that, depending on the antigens presented, leads to differentiation of these effector $CD4^+$ T-helper cells into subsets, the two main types of which are Th1

and Th2. Th1 cells are involved in cell-mediated immunity and phagocyte-dependent inflammation and produce IFN-γ, IL-2, and TNF-β (Romagnani 2000). Th2 cells, on the other hand, produce IL-4, IL-5, IL-6, IL-9, IL-10, and IL-13 and evoke strong antibody responses and eosinophil accumulation in addition to inhibiting phagocytic-cell functioning (Romagnani 2000). MiR-155 has a pivotal role in the regulation of both Th1 and Th2 cells. Overexpression of miR-155 in activated $CD4^+$ T cells promotes Th1 cell differentiation (Banerjee et al. 2010). Moreover, under Th1-inducing conditions, miR-155 targets IFN-γRα, which affects T-cell differentiation through cytokine signaling (Banerjee et al. 2010). On the other hand, suppression of miR-155 shifts the differentiation of $CD4^+$ T cells toward the Th2 phenotype (Chen et al. 2020) by increasing levels of the $CD4^+$ T-cell target c-Maf, which leads to enhanced production of the Th2 cytokines IL-4, IL-5, and IL-10 (Rodriguez et al. 2007). The role of miR-155 in adaptive immunity has significant oncogenic consequences; indeed, miR-155 expression has been directly associated with antitumor immune responses in 30 types of solid tumors, including melanoma, in humans (Huffaker et al. 2017).

Overexpression of the miR-17-92 cluster (consisting of miR-17, miR-18a, miR-19a, miR-19b, and miR-20a) in Th1 cells has been noted to have oncogenic effects in B cell lymphoma and prostate cancer (Kuo et al. 2019). Specifically, miR-17 and miR-19b regulate the Th1 response by promoting proliferation, reducing activation-induced cell death, enhancing IFN-γ production, and suppressing regulatory T-cell (Treg) differentiation. The regulatory functions of miR-17 and miR-19b result from their targeting TGF-βRII, CREB1, and PTEN (Jiang et al. 2011). Moreover, miR-146 has also been shown to inhibit STAT1 expression, thereby controlling Th1 response (Lu et al. 2010).

Another subset of $CD4^+$ T cells called follicular helper cells express chemokine receptor 5 and participate in the humoral immune response by regulating immune cell growth, differentiation, immunoglobulin isotype switching, affinity maturation of B cells, and antibody secretion, in part through their production of IL-21. The miR-17-92 cluster regulates the differentiation of follicular T-helper cells by targeting PTEN and the transcription factor RORα, which blocks the expression of genes associated with other T-helper cell subtypes such as Th17 and Th22 (Ranji et al. 2013). Th17 cells produce the cytokines IL-17A, IL17-F, and IL-22, and as such are also crucial in adaptive and innate immunity (Ranji et al. 2013). Th22 cells, the differentiation of which is promoted by miR-31, produce cytokine IL-22 (Jia and Wu 2014; Huang et al. 2019). The miR-17-92 cluster also regulates the migration of follicular T-helper cells by inhibiting PHLPP2, thereby accelerating P13K signaling and prompting migration of T cells into the B cell follicle and germinal center (Kuo et al. 2019).

Th17 cells can have both antitumor and pro-tumorigenic effects, the former through their ability to enhance $CD8^+$ T-cell function (Li et al. 2016) and the latter by inducing tumor vascularization and eliciting IL-6 production by both the tumor and tumor-associated stromal cells, leading to STAT3-mediated upregulation of pro-survival and pro-angiogenic genes (Wang et al. 2009). Mice deficient in IL-17A were found to be more likely to develop tumors, because Th17 cells activate

tumor-specific cytotoxic ($CD8^+$) T cells that are crucial to the antitumor effect (Kuen et al. 2020). In melanoma, the expression of IL-17 by Th17 cells promotes angiogenesis. The development of Th17 cells is regulated by two miRNAs, miR-326, and miR-181c. MiR-326 targets the negative regulator transcription factor ETs-1 for differentiation, and miR-181c targets the negative regulator Smad7, which results in increased TGF-β-induced Smad2/3 signaling and inhibition of Th17 cell differentiation via inhibition of IL-2 functions (Zhang et al. 2018c). MiR-10a also affects differentiation of Th17 and Treg cells from $CD4^+$ cells (Zhang et al. 2018c).

3.3.2 Cytotoxic T Cells

Recognition of antigens synthesized in the cellular cytoplasm prompts the differentiation of $CD8^+$ T cells into cytotoxic T cells, which in turn induces target cells to undergo programmed cell death. Overexpression of the miR cluster miR-17-92 by $CD8^+$ cells enhances IFN-γ production, ultimately driving cytotoxic T-cell responses that are promoted by Th1, and also leads to downregulation of TGF-βRII, which increases the frequency of memory T cells (Kosaka et al. 2015). Moreover, activation of the T-cell receptor results in upregulation of miR-155, which is necessary for $CD8^+$ cells to thrive and proliferate (Monnot et al. 2020), for limiting syngeneic tumor growth, and for promoting the production of IFN-γ by T cells (Huffaker et al. 2017). $CD8^+$ T cell proliferation is also regulated by miR-21, which targets dual-specificity phosphatase (DUSP10), and miR-30b, which targets B cell CLL/lymphoma 6 (BCL-6) (Chang et al. 2012). Transcription of both miR-21 and miR-155 relies on a negative feedback involving DUSP10, which downregulates the activity of the AP-1 necessary for transcribing these miRNAs (Chang et al. 2012). Downregulation of miR-21 by Spry2 has also been linked with malignant progression in gliomas (Kwak et al. 2011), thereby implicating Spry2 in the treatment of glioma resulting from posttranscriptional regulation by miR-21 (Kwak et al. 2011). miR-21 also regulates the polarization and inflammatory response of Th1 and Th2 cells through the IL-2 and IFN-γ pathways (Liu et al. 2014).

3.3.3 Regulatory T Cells

Tregs, a subset of $CD4^+$ T cells that express the surface antigen Foxp3, regulate immune function by limiting pathogenic immune responses to self-antigens and binding to the promotor region of the *BIC* gene, the sequence of which includes miR-155 (Eis et al. 2005). MiR-155 may also promote maturation of medullary thymic epithelial cells by contributing to thymic Treg cell differentiation (Dong et al. 2021). MiR-155 is upregulated in stimulated Tregs (Dong et al. 2021), and miR-155 deficiency is thought to dampen the development of Tregs, resulting in diminished numbers of both thymic and splenic Tregs (Kohlhaas et al. 2009). MiR-155 also

targets the negative regulator of 1 L-2R signaling SOCS1, thereby increasing the Th-17-mediated inflammatory response (Wang et al. 2018a). Expressed at low levels in Treg cells, miR-15 and miR-16 suppress the T cell cycle and memory T-cell differentiation; miR-15a and miR-16 can also target Foxp3 and CTLA4 (Liu et al. 2014). Overexpression of these miRNAs reduces Foxp3 and CTLA4 levels, thereby partially suppressing Tregs (Liu et al. 2014).

3.4 miRNA-Mediated Regulation in the Tumor Microenvironment

The TME is essential to tumor maintenance and progression, as it encompasses interactions among fibroblasts, endothelial cells, stromal cells, and immune cells (da Cunha et al. 2019). The TME has two roles in tumor development: first, it hosts a group of cells that is capable of identifying and eliminating tumor cells, which stops cancer progression; however, it is also home to growth factors and hormones that contribute to tumor cell progression (Hanahan and Weinberg 2000, 2011; Grivennikov et al. 2010; Beatty and Gladney 2015).

Immune regulatory cells such as MDSCs, Tregs, DCs, and TAMs in the TME produce numerous cytokines, interleukins, chemokines, and metabolites. Tumor-infiltrating immune cells (e.g., cytotoxic T cells and NK cells) also affect antitumor immunity through both tumor-suppressive and tumor-promoting activities (Giraldo et al. 2015; Joyce and Fearon 2015; Malekghasemi et al. 2020; Verneau et al. 2020; Wang et al. 2020b). From this perspective, miRNAs modulate the tumor immune response by regulating the recruitment of different immune cell sets in the TME; hence abnormal expression of miRNAs by tumor cells can affect the composition of the immune microenvironment. Clarifying the mechanisms by which tumor and immune cells interact would contribute substantially to the use of miRNAs as predictors of prognosis, diagnosis, and immunotherapy strategies.

The crosstalk between immune and tumor cells depends on the incorporation of molecules in circulation in the TME, among them miRNAs, which appear through the internalization of free miRNAs in extracellular space or by interactions with EVs such as apoptotic bodies, microvesicles, and exosomes (Kogure et al. 2019). The principal differences among these types of EVs are their origin and size. The apoptosis machinery produces apoptotic bodies that are 1–5 μm in diameter; microvesicles emerge from protuberances in the plasma membrane, and are 100–1000 nm in diameter (Heijnen et al. 1999; Turiák et al. 2011); and exosomes, the smallest of the EVs at 40–100 nm, are formed from the fusion of plasma and endosomal membranes (Heijnen et al. 1999; Belting and Wittrup 2008). EVs released from immune cells interact with tumors and vice versa, enabling the internalization of miRNAs to regulate gene expression and, consequently, stimulate biochemical mechanisms and evoke phenotypic changes in the recipient cells. This intercellular communication is particularly important because many miRNAs

regulate both immune and tumor cells (Anfossi et al. 2018). Studies of miRNAs over the past few decades have explored their potential use as biomarkers, mainly to evaluate the effectiveness of anticancer treatments (Anfossi et al. 2018). Indeed, levels of M2-like TAMs and the recruitment of MDSCs and Tregs can be linked to dysregulation of specific miRNAs in the TME.

TAMs, the most abundant of the immune cells in the TME, participate in tumor cell growth, inflammation, invasion, metastasis, angiogenesis, and immunoregulation. miRNAs participate in the regulation and polarization of TAMs (Malekghasemi et al. 2020). The polarization process contributes to tumor development, because M1-like macrophages are involved in the inflammatory response and antitumor immunity whereas M2-like macrophages have pro-tumorigenic and anti-inflammatory properties. Conversion of the pro-immune M1-like phenotype to an immune regulatory M2-like phenotype can result from aberrant miRNA regulation in tumors. The TME is also rich in soluble factors such as TGF-β, VEGF, IL-4, IL-13, IL-10, and PGE2 that promote macrophage polarization as well as having other pleiotropic immune-suppressive effects. Indeed, miR-125, miR-29, and miR-155 are well known for promoting the repolarization of macrophages from the M2 to the M1 phenotype (Essandoh et al. 2016; Wang et al. 2020a).

Conversely, some miRNAs can prevent repolarization of M2 macrophages; in one example, stimulation with LPS or IFN-γ leads to increases in miR-21, which can prevent PGE2-mediated M2 generation by targeting STAT3 (Wang et al. 2015b). In general, M1-like stimuli (that is, cytokines secreted by Th1 cells) leads to reductions in miR-23a, miR-27a, and miR-24-2 cluster expression by binding the NF-κB promoter. Conversely, M2-polarization stimuli (e.g., Th2 cytokines) can activate the expression of this cluster via STAT6 signaling (Ma et al. 2016). In colorectal cancer, cancer cell-derived EVs containing miR-203 and miR-145 can polarize macrophages to the M2 phenotype and contribute to distant metastasis (Shinohara et al. 2017; Takano et al. 2017). In lung cancer, miR-21 and miR-29 have been observed in tumor-EVs delivered to TAMs, resulting in activation of the NF-κB pathway by binding to TLR8 and TLR7 (Fabbri et al. 2012). Thus, miRNA transfer via tumor-EVs and cellular interactions in TME that contribute to macrophage reprogramming also contribute to tumor advancement.

Some miRNAs can regulate chemokines that lead to recruitment and infiltration of lymphocytes into tumors and the TME. MiR-21 contributes to lymphocyte migration; in one study, inhibition of this miR led to enhanced release of the chemokines CCL5 and CXCL10 in a breast cancer cell line, and consequently amplified lymphocyte migration (Wang et al. 2013b). In hepatocellular carcinoma, high levels of TGF-β suppressed the expression of miR-34a, which increased the production of the chemokine CCL2, which is responsible for recruiting Tregs to the TME (Yang et al. 2012). MiR-130a and miR-145 also target TGF-β receptors, and both are downregulated in myeloid cells. Ectopic expression of these miRs in myeloid cells promotes antitumor immunity by downregulating type 2 cytokines and increasing IFN-γ $CD8^+$ T cells, thereby restricting tumor metastasis (Ishii et al. 2018). In nasopharyngeal cancer, exosomal miR-24 reduces T cell proliferation and differentiation into Th1 and Th17 cells by increasing levels of pERK, pSTAT1, and

pSTAT3 and decreasing levels of pSTAT5 (Ye et al. 2016). The suppressor function of Tregs can be evoked by interaction with or by encompassing miRNAs in EVs. For example, miR-214 in tumor cell-EVs promotes the expansion of Tregs, enhancing immune suppression through PTEN as a target in $CD4^+$ T cells (Okoye et al. 2014). T cells with alterations in the expression of miR-9 and miR-181 as delivered by tumor-EVs trigger the apoptosis cascade and reduce proliferative mechanisms contributing to breast tumor progression (Jiang et al. 2020).

DCs in the TME can also promote tumor development via miRNAs. MiR-221 induces proliferation of hepatocellular carcinoma cells and inhibits the maturation of DCs (Fu et al. 2019). Abnormal expression of several miRNAs in DCs directly affects their maturation and function, such as miR-22, miR-128, miR-133a, miR146a, and miR-212-3p (Du et al. 2012; Ding et al. 2015; Liang et al. 2015; Gao et al. 2016). In an osteosarcoma model, overexpression of miR-133a suppressed the maturation and activation of DCs (Gao et al. 2016). MiR-22 and miR-128 have been shown to affect the function of DCs by targeting P38 (Liang et al. 2015). EVs released by pancreatic cancer cells affect the immune functions of DCs by inducing immune tolerance through the action of miR-212-3p (Ding et al. 2015). EV-mediated transfer of miR-203 from pancreatic cancer cells to DCs induced downregulation of TLR4, TNF-α, and IL-12, which collectively suppressed activation of the immune response (Zhou et al. 2014).

Tumor EV miRNAs can also regulate some NK cell activities. As noted previously, NK cells are responsible for identifying and eliminating tumor cells by generating cytotoxic molecules such as perforin and granzymes. In one study of lung cancer cells, EVs originating from hypoxic tumor cells had elevated levels of miR-23a and miR-210 expression, which affected NK cytotoxicity (Berchem et al. 2016). In cervical cancer, overexpression of miR-20a in NK cells inhibited their antitumor effects by targeting RUNX1 (Zhu et al. 2018); overexpression of miR-24 in colorectal cancer cells and miR-218-5p in lung adenocarcinoma cells had similar effects (Zhang et al. 2018a; Yang et al. 2019).

MiRNAs also participate in the regulation of function, differentiation, and maturation of MDSCs. In breast cancer, miR-9 and miR-181a not only influence T cells in the TME but also affect the early development of MDSCs via interactions with SOCS3 and PIAS3 (Jiang et al. 2020). MiR-107 has been shown to interact with PTEN and DICER in MDSCs in gastric tumors (Ren et al. 2019b). Abnormal levels of some miRNAs influence the aggregation and other behavior of MDSCs in the TME, such as miR-30a, miR-155, miR-223, and miR-486 (Liu et al. 2011; Tian et al. 2015; Wang et al. 2015a; Xu et al. 2017b; Jiang et al. 2018).

In summary, strategies to increase the activity of the immune system require further exploration of how miRNAs interact with macrophages, DCs, NK cells, T-helper cells, cytotoxic T cells, MDSCs, and Tregs in tumors and the TME to participate in tumor development.

3.5 Immune Checkpoint Molecules

Immune checkpoint molecules are key regulators of the immune system. Two signals are required for a T cell to respond: the first is the interaction of major histocompatibility complex (MHC) proteins expressed by antigen-presenting cells with T-cell receptors, which confers the specificity of T-cell activation; the second signal involves co-stimulatory and co-inhibitory molecules that can determine whether the T cell is activated or suppressed. Those co-signaling molecules are known as immune checkpoints (Pardoll 2012). The best studied of these molecules involved in T-cell activation are PD1, expressed on T cells, and its ligand PDL1, expressed on various immune and nonimmune cell types. The interaction of those two molecules results in inhibition of T-cell activation and proliferation, as well as downregulation of pro-inflammatory cytokines and antiapoptotic proteins (Dermani et al. 2019).

Another widely investigated set of receptors and ligands is CTLA4 and CD80 or CD86. Interactions between the CTLA4 receptor, expressed mainly on T cells, and its ligands, expressed on several types of cells, lead to suppression of T-cell activity or promotion of Treg development and function (Wing et al. 2008; Hosseini et al. 2020). In the context of cancer, tumor cells can overexpress these inhibitory molecules to escape from immune surveillance.

MiRNAs are important regulators of immune checkpoint molecule expression, and the identification of relevant miRNAs could be beneficial to improve alternatives to immunotherapy (Yang et al. 2018). For example, PD1 can be directly regulated by binding between miR-138 (Wei et al. 2016), miR-149-3p (Zhang et al. 2019), and miR-374b (Huang et al. 2018) and the 3′-UTR of its mRNA (Table 3.1). PDL1 can also be directly regulated by several miRNAs, including miR-34a (Anastasiadou et al. 2019), miR-138 (Zhao et al. 2016; Rasoolnezhad et al. 2021), miR-142 (Berrien-Elliott et al. 2019), miR-155 (Yee et al. 2017), miR-200 (Chen et al. 2014; Katakura et al. 2020), miR-424 (Xu et al. 2016), miR-513 (Gong et al. 2010), and miR-570 (Wang et al. 2012) (Table 3.1). MiR-138 is also associated with the regulation of CTLA4 mRNA (Wei et al. 2016) (Table 3.1).

The CD28 receptor is a CTLA4 homolog and competes for the same ligands. However, these two molecules have opposite functions in T-cell activation: CTLA4 has an inhibitory role in the adaptive immune response and CD28 is a co-stimulatory molecule, enhancing T-cell survival through activation of AP-1 and NF-κB (Alegre et al. 2001; Rowshanravan et al. 2018). MiR-145 (Wang et al. 2013a) and the miR-17-92 cluster (de Kouchkovsky et al. 2013) can interact with CD28. Because both receptors (CTLA4 and CD28) bind to the same ligands, dysregulation of those miRNAs could lead to lower levels of CD28, resulting in a more suppressive profile.

Another immune checkpoint molecule with bidirectional roles in immune activation is B7-H3 (also known as CD276), which can act as a co-stimulatory or co-inhibitory molecule. The receptor that binds to B7-H3 is still not confirmed; however, the most probable candidate is TLT-2 (Yang et al. 2020). B7-H3 also has non-immunologic functions in cancer that lead to proliferation, migration, invasion, angiogenesis, and drug resistance (Flem-Karlsen et al. 2020). Several miRNAs can

target the 3′-UTR of B7H3 mRNA, including miR-29c, the overexpression of which in breast cancer patients was associated with a reduced risk of death by downregulating B7-H3 (Nygren et al. 2014).

In addition to direct regulation, miRNAs can regulate immune checkpoint molecules by regulating their upstream signaling pathways. The Epstein-Barr virus miRNA, miR-BART-5p, is one example, as it can directly target PIAS3 mRNA, downregulating PIAS3 protein levels with consequent activation of STAT3 and upregulation of PDL1 (Yoon et al. 2020).

A novel class of repetitive short DNA sequences, known as pyknons, are transcribed and present in nearly all mRNAs, and can contain binding sites for miRNAs (Rigoutsos et al. 2006). Pyknon-90, which is present in the sequence of the long noncoding RNA N-BLR, harbors a binding site for miR-200c-3p (Rigoutsos et al. 2017). This primate-specific long noncoding RNA can act as a "sponge" for miR-200, increasing the levels of its targets such as PDL1. This example illustrates another layer of regulatory mechanisms between miRNAs and immune checkpoint molecules.

In summary, identifying relevant miRNAs that can directly or indirectly affect immune checkpoint inhibitors that are overexpressed in cancer cells could lead to additional alternatives to antitumor immunotherapies, and perhaps predict responsiveness to treatments.

3.6 miRNAs as Potential Targets for Immunotherapy

As alluded to earlier in this chapter, miRNAs regulate the expression of several gene targets that themselves are key regulators of various signaling pathways, including immune signaling (Di Martino et al. 2021; Raue et al. 2021). For this reason, increased interest has been expressed in exploring the potential role of miRNAs as therapeutic agents in cancer immunotherapy. Depending on their function, miRNAs can be oncogenic ("oncomiRs") or tumor suppressive and are commonly dysregulated in several types of cancer (Table 3.2). The twin goals of miRNA-based therapy are to restore or replace downregulated tumor suppressor miRNAs and to inhibit upregulated oncomiRs (Mollaei et al. 2019; Di Martino et al. 2021; Raue et al. 2021).

One way of restoring the expression of tumor suppressor miRNAs is by using miRNA mimics, synthetic double-strand oligonucleotides with the same sequence as the mature target miRNA (Mollaei et al. 2019; Di Martino et al. 2021; Raue et al. 2021). However, use of miRNA mimics as a therapeutic strategy presents some difficulties related to their stability and delivery. To enable systemic delivery and increase the stability of miRNA mimics, several chemical modifications have been tested, such as including 2′-OH modifications and the use of peptide nucleic acids and locked nucleic acids; conjugation with small cell peptides and aptamers has also been used for delivery purposes. Restoration of stably expressed specific miRNAs can also be achieved with viral vectors, such as lentivirus, adenovirus, and adeno-associated viruses; other nonviral delivery methods involve inorganic, lipid-based, and polymeric carriers (Mollaei et al. 2019; Di Martino et al. 2021; Raue et al. 2021).

Table 3.2 MiRNAs that are potential targets in cancer immunotherapy

miRNA	Action (Mimic or Inhibitor)	Immune response	Types of cancer studied	Clinical trials	References
miR-100	Antagomir	Induces M2 polarization	Breast cancer	–	Wang et al. (2018d)
miR-125b	Mimic	Enhances apoptosis and pro-inflammatory status; enhances iNOS/Arg1 ratio, TNF-α, and IL-1; reduces IL-10; induces M1 polarization	Lung adenocarcinoma, hepatocellular carcinoma	–	Talekar et al. (2016), Wang et al. (2020a)
miR-128	Mimic	Suppresses tumor growth and metastasis; increases numbers of DCs, $CD8^+$ T cells, and NK cells in tumors; promotes antitumor immunity	Pancreatic adenocarcinoma	–	Xi et al. (2020)
miR-138	Mimic	Reduces levels of immune checkpoints CTLA4, PDL1, and FoxP3	Glioma	–	Wei et al. (2016)
miR-142-5p	Mimic	Inhibits PDL1 expression and increases $CD4^+$ and $CD8^+$ T cells	Breast cancer	–	Jia et al. (2017)
miR-143	Mimic	Enhances the cytotoxicity of CAR-T cells	Esophageal cancer	–	Zhang et al. (2018b)
miR-146a-5p	AntagomiR	Reduces tumor growth, angiogenesis, and TGF-β levels; increases apoptosis and $CD8^+$ T cells	Renal cortical adenocarcinoma	–	Simanovich et al. (2018)
miR-155	AntagomiR	Related to different immunologic processes	Hematologic cancers, cutaneous T cell lymphomas	NCT02580552 NCT03713320	Ji et al. (2015), Mashima (2015), Foss et al. (2018)
miR-16	Mimic	Related to immune checkpoints; reduces PDL1; promotes T cell response; decreases tumor growth	Pleural mesothelioma	NTC02369198	van Zandwijk et al. (2017), Li et al. (2020)
miR-17-92	Mimic	Improves survival and cytotoxic activity	Glioblastoma	–	Ohno et al. (2013)
miR-203	Mimic	Reduces TLR4, TNF-α, and IL-12 levels; reduces antitumor immunity	Pancreatic adenocarcinoma	–	Zhou et al. (2014)

(continued)

Table 3.2 (continued)

miRNA	Action (Mimic or Inhibitor)	Immune response	Types of cancer studied	Clinical trials	References
miR-23a-3p	Mimic	Upregulates PDL1 expression; stimulates immune escape via the PTEN-AKT/PI3K pathway	Hepatocellular carcinoma	–	Liu et al. (2019)
miR-27a-3p	Mimic	Upregulates PDL1 expression; stimulates immune escape via the PTEN-AKT/PI3K pathway	Breast cancer	–	Yao et al. (2020)
miR-34	Mimic	Promotes the antitumor immune response	[a]	NCT01829971	Bader (2012), Cortez et al. (2016)

Abbreviations: DCs, dendritic cells; NK cells, natural killer cells; CAR, chimeric antigen receptor

[a]Includes hepatocellular carcinoma, melanoma, renal cell carcinoma, lung cancer, gastrointestinal stromal tumor, neuroendocrine tumors, and other tumor types

The activity of oncogenic miRNAs can be inhibited either directly by blocking binding of the miRNA to its target, or indirectly by interacting with the target so as to prevent miRNA binding (Mollaei et al. 2019; Di Martino et al. 2021; Raue et al. 2021). Molecules used for this purpose include synthetic antisense oligonucleotides (ASOs), locked nucleic acid antisense oligonucleotides, miRNA "sponges," miRNA masks, and nanoparticles. ASOs, single-stranded nucleic acids complementary to a mature miRNA target, are the most commonly used miRNA inhibitors. ASOs can be chemically modified to enhance their stability and efficiency of delivery. The miRNA "sponges" function as decoys that prevent the miRNA from binding to its target. These transcripts can be targeted to a specific miRNA or to a set of miRNAs that share the same binding site. These two approaches are based on complementary sequences in mature miRNA. The miR-masks, in contrast, act by masking the binding site on the mRNA target, thereby precluding the miRNA–mRNA interaction (Mollaei et al. 2019; Di Martino et al. 2021; Raue et al. 2021).

TAMs, the central immune cell type in the TME, can be tumor-promoting or tumor-inhibiting, depending on their phenotype (M0, M1, or M2). In one in vivo study of KP (KrasG12D/p53fl/fl) mice, human lung adenocarcinoma cells were transfected with miR125b- and wt-p53-expressing plasmids in dual D44/EGFR-targeted hyaluronic acid-based nanoparticles, which led to enhanced apoptosis and pro-inflammatory status, as characterized by increased iNOS/Arg1 ratio, TNF-α, and IL-1β and reduced IL-10 levels. In the same study, coculturing J774.A1 macrophages with the transfected SK-LU-1 lung adenocarcinoma cells in vitro led to repolarization of the macrophages (Talekar et al. 2016). In another study with mouse models of hepatocellular carcinoma and Lewis lung cancer, the targeted delivery of miR-125a or miR-99b into TAMs led to repolarization of M2-like TAMs toward the M1 phenotype (by downregulating κB-Ras2) and reduced the M2 phenotype (by downregulating mTOR), leading to tumor regression (Wang et al. 2020a).

In addition to miRNA restoration therapy, inhibition of miRNAs has also been used to repolarize macrophages. In an example of this approach, use of an miR-100 antagomiR-induced M2 polarization in RAW 264.7 cells and mouse peritoneal macrophages cultured in vitro. Moreover, M2 TAMs were also observed after the intratumoral injection of the miR-100 antagomiR in a 4 T1 mouse breast cancer model (Wang et al. 2018d).

In addition to affecting macrophage polarization, miRNAs also regulate the activity of other immune cells and key receptors in antitumor responses. In one study of the anticancer potential of an miR-146a antagomir, its systemic injection, in combination with intratumoral injection of stimulated RAW 264.7 macrophages, led to reduced tumor growth and angiogenesis, and increased apoptosis, in BALB c/mice implanted with RENCA mouse renal carcinoma cells. This combined approach also affected the TME by increasing $CD8^+$ T cells and reducing levels of the M2-related cytokine TGF-β (Simanovich et al. 2018). Another approach involved investigating the influence of another miRNA, miR-203, on TLR expression and downstream cytokines in DCs. In that study, culturing DCs with pancreatic cancer-derived exosomes (which had high miR-203 levels) and miR-203 mimics led

to decreased levels of TLR4, TNF-α and IL-12; however, treatment with exosomes and miR-203 inhibitors reversed this downregulation. Thus the authors concluded TLR4 has an essential role in the maturation of DCs and is regulated by miR-203 (Zhou et al. 2014).

Another miRNA with great immunotherapeutic potential is miR-128, which regulates the activity of several types of immune cells in the TME and can suppress tumor growth and metastasis in a pancreatic adenocarcinoma mouse model. Indeed, injection of an miR-128 overexpressing lentivirus vector enhanced the numbers of DCs, $CD8^+$ T cells, and NK cells in both the tumor and the spleen in that mouse model, and promoted antitumoral immunity via the ZEB1/CD47 axis (Xi et al. 2020). Blockade of CD47 has also emerged as another promising form of tumor immunotherapy (Matlung et al. 2017; Weiskopf 2017; Murata et al. 2018). CD47 mediates immune evasion by its interaction with signal regulatory protein (SIRP)-α expressed on macrophages and other myeloid cells, which leads to inhibition of phagocytosis; thus CD47-blocking approaches would restore phagocytosis and enhance antitumor effects (Matlung et al. 2017; Weiskopf 2017; Murata et al. 2018).

As noted previously, exosomes, which carry several regulatory molecules including miRNAs, are also being tested as anticancer therapy. One group used exosome cocultures with macrophages in vitro as well as injecting exosomes in vivo to demonstrate that the upregulation of PDL1 expression in macrophages is promoted by miR-23a-3p (Liu et al. 2019). Specifically, endoplasmic reticulum-stressed hepatocellular carcinoma cells release exosomes with abundant levels of miR-23a-3p, which inhibits PTEN expression and enhances phosphorylated AKT and PDL1 expression in macrophages. In another study, exosomes derived from endoplasmic reticulum-stressed breast cancer cells showed high levels of miR-27a-3p and induced PDL1 expression in macrophages in vitro and in vivo via the PTEN-AKT/PI3K pathway, with the ultimate result of stimulating immune escape (Yao et al. 2020).

Another miRNA that seems to participate in PDL1 regulation is miR-142-5p. Overexpression of this miRNA inhibits PDL1 expression, increases the numbers of $CD4^+$ and $CD8^+$ T cells, and reduces the numbers of $PD1^+$ T cells in vivo, thereby enhancing antitumoral immunity (Jia et al. 2017). Another miRNA, miR-138, regulates the expression of another immune checkpoint molecule, CTLA4, in addition to PDL1. In one study, transfection of human $CD4^+$ T cells with miR-138 led to reduced levels of CTLA4, PDL1, and Foxp3 (a marker of Tregs); giving miR-138 to immune-competent mice implanted with GL261 glioma cells had the same effect (Wei et al. 2016).

MiRNAs have also been combined with CAR-T cells with the goal of enhancing their cytotoxic effects. In one such study, CAR-T cells directed to glioblastoma cells stably expressing EGFR variant III co-transduced with miR-17-92 led to improved survival and cytotoxic activity under temozolomide treatment (Ohno et al. 2013). Another example of a synergistic effect between CAR-T cell therapy and miRNA expression was noted in the esophageal cancer cell line TE7; in that study, the cytotoxicity of HER2-CAR-T cells against TE-7 cells was enhanced by miR-143 overexpression (Zhang et al. 2018b).

Despite these exciting findings, few miRNAs have been evaluated to date in clinical trials of miRNA restoration or miRNA inhibition for cancer therapy. The first miRNA tested for miRNA-based anticancer therapy was miR-34, which is known to be downregulated in several types of cancer. In trial NCT01829971, a liposomal mimic of miR-34a (MRX34; Mirna Therapeutics Inc.), was given by injection in patients with solid tumors refractory to standard treatments (e.g., hepatocellular carcinoma, melanoma, renal cell carcinoma, lung cancer, gastrointestinal stromal tumor, and neuroendocrine tumors), with the goal of restoring miR-34 levels and its activity in the p53/WNT signaling pathway. The challenges involved in this therapeutic approach included the appearance of immune-related adverse events in the phase I portion of the trial. In the subsequent trial NCT02862145, a recommended phase II dose of MRX34 was evaluated in patients with advanced melanoma; that dose, preceded by dexamethasone, had an acceptable toxicity profile for most patients (Hong et al. 2020). In a preclinical study, MRX34 was found to increase the tumor infiltration of $CD8^+$ T cells and to reduce the number of $CD8^+PD1^+$ T cells in a syngeneic mouse model of non-small cell lung cancer; the addition of radiation therapy was further enhanced the number of $CD8^+$ T cells (Cortez et al. 2016). These findings suggest that miR-34a mimics have potential for promoting an antitumor immune response through its actions as a master tumor suppressor that targets a variety of oncogenic genes, thereby having effects on several signaling pathways (Bader 2012).

In addition to restoration therapy, another approach involves use of a locked nucleic acid-based antagonist of miR-155, MRG-106 (cobomarsen), which is being evaluated in a phase I study of patients with hematologic malignancies including chronic lymphocytic leukemia (NCT02580552) and a phase II trial of patients with cutaneous T cell lymphoma (NCT03713320) (Foss et al. 2018). As noted elsewhere in this chapter, miR-155 has many effects on a diverse variety of immune cells; the reasoning behind using miR-155 antagonists is that its deficiency impairs immune functions (Ji et al. 2015; Mashima 2015).

In another approach to miRNA therapy, miR-16 mimics have been assessed in therapy for pleural mesothelioma (NTC02369198). This novel approach involves use of "TargomiRs" that is, minicells (EnGeneIC Dream Vector) loaded with a miR-16 mimic. Intravenous injection of TargomiRs in a phase I study revealed an acceptable safety profile (van Zandwijk et al. 2017). Finally, the mechanism of action of the related molecule miR-16a-5p was recently found to be linked with the immune checkpoint molecule PDL1 in gastric cancer. That study revealed that miR-16-5p was transferred between M1 macrophages and gastric cancer cells via exosomes and targeted PDL1, and that M1 macrophage-derived exosomes containing miR-16-5p triggered a T-cell immune response leading to inhibited tumor formation, both in vitro and in vivo, by reducing the expression of PDL1 (Li et al. 2020).

MiRNAs that could serve as targets in cancer immunotherapy are summarized in Table 3.2.

3.7 Challenges in miRNA-Based Therapy

Although miRNAs have potential as immunotherapy or as targets in immunotherapy, their implementation in clinical practice is still far away. Among the various strategies tested to date, the combination of miRNA-based therapy with other methods, e.g., chemotherapy (Chakraborty et al. 2018) or silencing with siRNAs, had superior results, such as enhancing antitumor effectiveness (Nishimura et al. 2013). However, estimating the correct dose of each component in such combined therapies can be a challenge (Yin et al. 2018) and must be customized for each patient. Three considerations regarding miRNA-based therapy are the need for miRNA stability in vivo; tumor-specific delivery; and retention in metabolizing organs such as liver and kidney, although these points could be addressed by delivering miRNAs via nanoparticles. Combined therapies could help to develop synergistic effects (Chakraborty et al. 2018).

In one example, a miR-5119 mimic-engineered DC vaccine was able to increase antitumor immune response in breast cancer cells by increasing cytokine production and reducing T-cell apoptosis, as well as reducing tumor size and volume. However, although DCs can increase the activity of T cells, DCs can also express ligands that bind to T cells (e.g., PDL1) that induce T cell exhaustion (Zhang et al. 2020).

In general, studies of miRNAs in immunotherapy done to date have some limitations, including small numbers of patients and lack of standardization with regard to documentation of clinicopathological factors, habits, and treatment that could affect the upregulation or downregulation of miRNAs (Peng et al. 2020). miRNAs can also affect the TME in ways that lead to decreased cell migration; in one study, for example, miR-1 contributed to resistance to EGFR-TKI immunotherapy and prompted the induction of cytokines, leading the authors to propose that miR-1 could be used as indicator of the effectiveness of chemotherapy as well (Kawana et al. 2021).

Exosomes, because they are endogenous molecules, may be one of the best options for therapies that involve delivery of miRNAs, but several limitations to this approach must also be acknowledged. Difficulties in producing and isolating exosomes have led to generally low yields; strategies such as the use of bioreactors or modulating oxygen or pH levels could help to bring production to a larger scale. However, these approaches must be considered carefully, as stress can affect the contents of exosomes. Also, the technology for generating exosomes is not well developed, and difficulties with storage can affect the purity of the product. Finally, the optimal route of administration of miRNAs by various delivery methods must also consider their biodistribution (Yamashita et al. 2018).

3.8 Conclusions

MiRNAs are deeply involved in the regulation of numerous anticancer immune responses (Fig. 3.1), including both innate and adaptive immunity. Although miRNA-based therapy has yet to become a reality in clinical settings, the use of

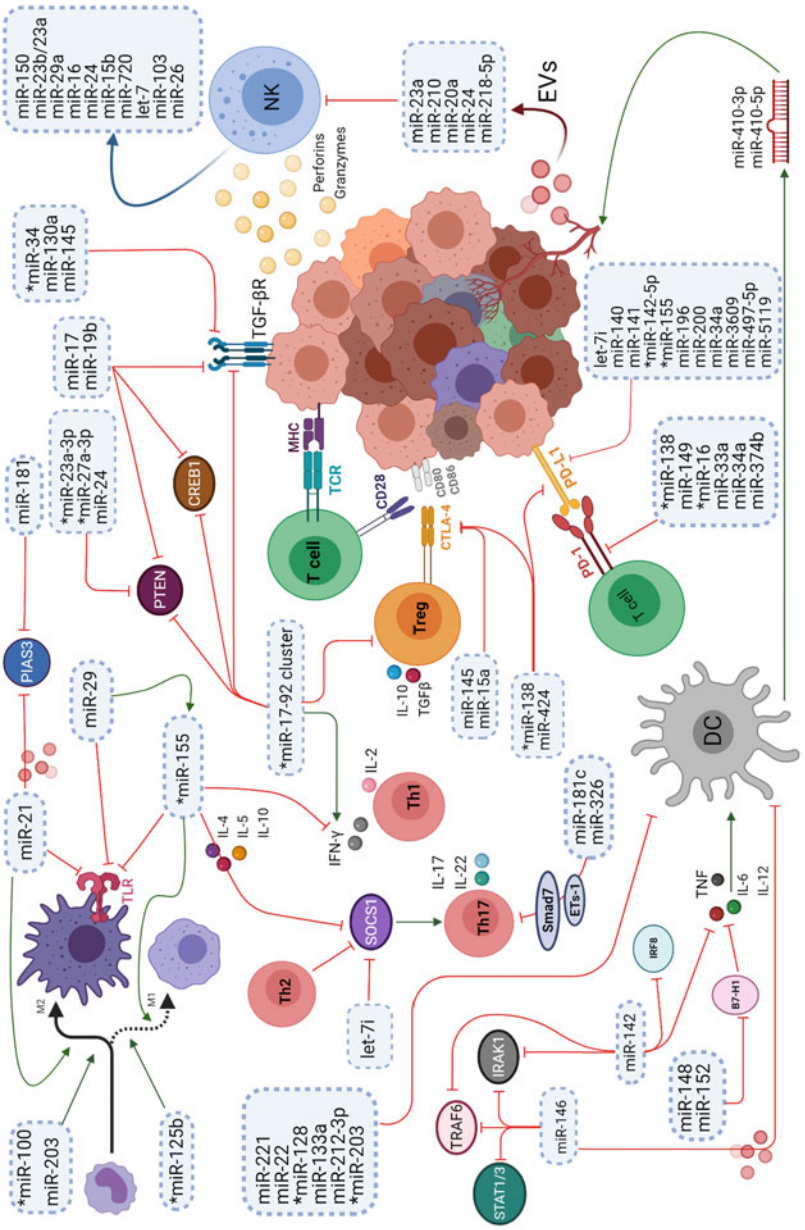


Fig. 3.1 Interactions among immune cells and miRNAs in the tumor microenvironment. *miRNAs that have been studied as potential targets for immunotherapy. Abbreviations: CD80L, cluster of differentiation 80 ligand; CD86L, cluster of differentiation 86 ligand; CREB1, CAMP-responsive element binding protein 1; CTLA4, cytotoxic T-lymphocyte associated protein 4; DCs, dendritic cells; EVs, extracellular vesicles; IL-6, interleukin-6; IL-10, interleukin-

miRNA mimics and inhibitors as immunotherapeutics has huge potential. However, some challenges remain that must be overcome, such as identifying which patients would truly benefit from miRNA-based therapy; reducing immune-related adverse events; and improving drug delivery strategies. Thus, much remains to be determined regarding how miRNAs interact with aspects of the TME to fully recognize the potential of miRNAs in immunotherapy against cancer.

Acknowledgments This work was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) as well as Cancer Center Support (Core) Grant P30-CA016672 from the National Cancer Institute, US National Institutes of Health, and Department of Defense (DOD) Lung Cancer Research Program (LCRP) W81XWH-21-1-0336 to The University of Texas MD Anderson Cancer Center.

References

Alegre ML, Frauwirth KA, Thompson CB (2001) T-cell regulation by CD28 and CTLA-4. Nat Rev Immunol 1:220–228. https://doi.org/10.1038/35105024

Anastasiadou E, Stroopinsky D, Alimperti S et al (2019) Epstein−Barr virus-encoded EBNA2 alters immune checkpoint PD-L1 expression by downregulating miR-34a in B-cell lymphomas. Leukemia 33:132–147. https://doi.org/10.1038/s41375-018-0178-x

Anfossi S, Fu X, Nagvekar R, Calin GA (2018) MicroRNAs, regulatory messengers inside and outside cancer cells. Adv Exp Med Biol 1056:87–108. https://doi.org/10.1007/978-3-319-74470-4_6

Angius A, Uva P, Pira G et al (2019) Integrated analysis of miRNA and mRNA endorses a twenty miRNAs signature for colorectal carcinoma. Int J Mol Sci 20:1–16. https://doi.org/10.3390/ijms20164067

Bader AG (2012) miR-34 - a microRNA replacement therapy is headed to the clinic. Front Genet 3:120. https://doi.org/10.3389/fgene.2012.00120

Banerjee A, Schambach F, Dejong CS et al (2010) Micro-RNA-155 inhibits IFN-γ signaling in CD4+ T cells. Eur J Immunol 40:225–231. https://doi.org/10.1002/eji.200939381

Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136:215–233. https://doi.org/10.1016/j.cell.2009.01.002

Beatty GL, Gladney WL (2015) Immune escape mechanisms as a guide for cancer immunotherapy. Clin Cancer Res 21:687–692. https://doi.org/10.1158/1078-0432.CCR-14-1860

Belting M, Wittrup A (2008) Nanotubes, exosomes, and nucleic acid-binding peptides provide novel mechanisms of intercellular communication in eukaryotic cells: implications in health and disease. J Cell Biol 183:1187–1191. https://doi.org/10.1083/jcb.200810038

Belz GT (2013) miR-142 keeps CD4+ DCs in balance. Blood 121:871–872. https://doi.org/10.1182/blood-2012-12-472589

Fig. 3.1 (continued) 10; IL-12, interleukin-12; INF-γ, interferon-gamma; IRAK1, interleukin-1 receptor-associated kinase 1; MHC, major histocompatibility complex; NK, natural killer cells; PIAS3, protein inhibitor of activated STAT3; PD1, programmed cell death protein 1; PDL1: programmed death-ligand 1; PTEN, phosphatase and tensin homolog; SOCS1, suppressor of cytokine signaling 1; STAT, signal transducer and activator of transcription; TGF-BR, transforming growth factor beta receptor; TLR, Toll-like receptor; TRAF6, TNF receptor-associated factor 6; Treg, regulatory T cell

Berchem G, Noman MZ, Bosseler M et al (2016) Hypoxic tumor-derived microvesicles negatively regulate NK cell function by a mechanism involving TGF-β and miR23a transfer. Onco Targets Ther 5:1–13. https://doi.org/10.1080/2162402X.2015.1062968

Berrien-Elliott MM, Sun Y, Neal C et al (2019) MicroRNA-142 is critical for the homeostasis and function of type 1 innate lymphoid cells. Immunity 51:479–490.e6. https://doi.org/10.1016/j.immuni.2019.06.016

Bezman NA, Cedars E, Steiner DF et al (2010) Distinct requirements of microRNAs in NK cell activation, survival, and function. J Immunol 185:3835–3846. https://doi.org/10.4049/jimmunol.1000980

Cai X, Yin Y, Li N et al (2012) Letter to the editor re-polarization of tumor-associated macrophages to pro-inflammatory M 1 macrophages. J Mol Cell Biol 4:1–3. https://doi.org/10.1093/jmcb/mjs044

Chakraborty C, Sharma AR, Sharma G et al (2018) The novel strategies for next-generation cancer treatment: miRNA combined with chemotherapeutic agents for the treatment of cancer. Oncotarget 9:10164–10174. https://doi.org/10.18632/oncotarget.24309

Chang C-C, Zhang Q-Y, Liu Z et al (2012) Downregulation of inflammatory microRNAs by Ig-like transcript 3 is essential for the differentiation of human CD8 + T suppressor cells. J Immunol 188:3042–3052. https://doi.org/10.4049/jimmunol.1102899

Chen T, Li Z, Jing T et al (2011) MicroRNA-146a regulates the maturation process and pro-inflammatory cytokine secretion by targeting CD40L in oxLDL-stimulated dendritic cells. FEBS Lett 585:567–573. https://doi.org/10.1016/j.febslet.2011.01.010

Chen L, Gibbons DL, Goswami S et al (2014) Metastasis is regulated via microRNA-200/ZEB1 axis control of tumour cell PD-L1 expression and intratumoral immunosuppression. Nat Commun 5. https://doi.org/10.1038/ncomms6241

Chen L, Gao D, Shao Z et al (2020) miR-155 indicates the fate of CD4+ T cells. Immunol Lett 224:40–49. https://doi.org/10.1016/j.imlet.2020.05.003

Chow MT, Möller A, Smyth MJ (2012) Inflammation and immune surveillance in cancer. Semin Cancer Biol 22:23–32. https://doi.org/10.1016/j.semcancer.2011.12.004

Cortez MA, Ivan C, Valdecanas D et al (2016) PDL1 regulation by p53 via miR-34. J Natl Cancer Inst 108. https://doi.org/10.1093/jnci/djv303

Croce CM, Calin GA (2005) miRNAs, cancer, and stem cell division. Cell 122:6–7. https://doi.org/10.1016/j.cell.2005.06.036

Cui ZJ, Xie XL, Qi W et al (2019) Cell-free miR-17-5p as a diagnostic biomarker for gastric cancer inhibits dendritic cell maturation. Onco Targets Ther 12:2661–2675. https://doi.org/10.2147/OTT.S197682

Curtale G (2018) MiRNAs at the crossroads between innate immunity and cancer: focus on macrophages. Cell 7:12. https://doi.org/10.3390/cells7020012

Curtale G, Mirolo M, Renzi TA et al (2013) Negative regulation of toll-like receptor 4 signaling by IL-10-dependent microRNA-146b. Proc Natl Acad Sci U S A 110:11499–11504. https://doi.org/10.1073/pnas.1219852110

da Cunha BR, Domingos C, Buzzo Stefanini AC et al (2019) Cellular interactions in the tumor microenvironment: the role of secretome. J Cancer 10:4574–4587. https://doi.org/10.7150/jca.21780

Dahiya DS, Wani F, Guidi JC, Kichloo A (2020) Gastrointestinal adverse effects of immunotherapeutic agents: a systematic review. Gastroenterol Res 13:227–232. https://doi.org/10.14740/gr1340

de Kouchkovsky D, Esensten JH, Rosenthal WL et al (2013) microRNA-17–92 regulates IL-10 production by regulatory T cells and control of experimental autoimmune encephalomyelitis. J Immunol 191:1594–1605. https://doi.org/10.4049/jimmunol.1203567

De Martin E, Michot JM, Rosmorduc O et al (2020) Liver toxicity as a limiting factor to the increasing use of immune checkpoint inhibitors. JHEP Rep 2:100170. https://doi.org/10.1016/j.jhepr.2020.100170

Dermani FK, Samadi P, Rahmani G et al (2019) PD-1/PD-L1 immune checkpoint: potential target for cancer therapy. J Cell Physiol 234:1313–1325. https://doi.org/10.1002/jcp.27172

Di Martino MT, Riillo C, Scionti F et al (2021) Mirnas and lncrnas as novel therapeutic targets to improve cancer immunotherapy. Cancers (Basel) 13:1–27. https://doi.org/10.3390/cancers13071587

Ding G, Zhou L, Qian Y et al (2015) Pancreatic cancer-derived exosomes transfer miRNAs to dendritic cells and inhibit RFXAP expression via miR-212-3p. Oncotarget 6:29877–29888. https://doi.org/10.18632/oncotarget.4924

Donatelli SS, Zhou JM, Gilvary DL et al (2014) TGF-β-inducible microRNA-183 silences tumor-associated natural killer cells. Proc Natl Acad Sci U S A 111:4203–4208. https://doi.org/10.1073/pnas.1319269111

Dong J, Warner LM, Lin LL et al (2021) miR-155 promotes T reg cell development by safeguarding medullary thymic epithelial cell maturation. J Exp Med 218:e20192423. https://doi.org/10.1084/jem.20192423

Drake CG, Jaffee E, Pardoll DM (2006) Mechanisms of immune evasion by tumors. Adv Immunol 90:51–81. https://doi.org/10.1016/S0065-2776(06)90002-9

Du J, Wang J, Tan G et al (2012) Aberrant elevated microRNA-146a in dendritic cells (DC) induced by human pancreatic cancer cell line BxPC-3-conditioned medium inhibits DC maturation and activation. Med Oncol 29:2814–2823. https://doi.org/10.1007/s12032-012-0175-2

Dunand-Sauthier I, Santiago-Raber ML, Capponi L et al (2011) Silencing of c-Fos expression by microRNA-155 is critical for dendritic cell maturation and function. Blood 117:4490–4500. https://doi.org/10.1182/blood-2010-09-308064

Eis PS, Tam W, Sun L et al (2005) Accumulation of miR-155 and BIC RNA in human B cell lymphomas. Proc Natl Acad Sci U S A 102:3627–3632. https://doi.org/10.1073/pnas.0500613102

El Gazzar M, Church A, Liu T, McCall CE (2011) MicroRNA-146a regulates both transcription silencing and translation disruption of TNF-α during TLR4-induced gene reprogramming. J Leukoc Biol 90:509–519. https://doi.org/10.1189/jlb.0211074

Espinoza LJ, Takami A, Yoshioka K et al (2012) Human microRNA-1245 down-regulates the NKG2D receptor in natural killer cells and impairs NKG2D-mediated functions. Haematologica 97:1295–1303. https://doi.org/10.3324/haematol.2011.058529

Essandoh K, Li Y, Huo J, Fan G-C (2016) MiRNA-mediated macrophage polarization and its potential role in the regulation of inflammatory response. Shock 46:122–131. https://doi.org/10.1097/SHK.0000000000000604

Fabbri M (2020) Natural killer cell–derived vesicular miRNAs: a new anticancer approach? Cancer Res 80:17–22. https://doi.org/10.1158/0008-5472.CAN-19-1450

Fabbri M, Paone A, Calore F et al (2012) MicroRNAs bind to toll-like receptors to induce prometastatic inflammatory response. Proc Natl Acad Sci U S A 109:1–7. https://doi.org/10.1073/pnas.1209414109

Farazi TA, Spitzer JI, Morozov P, Tuschl T (2011) miRNAs in human cancer. J Pathol 223:102–115. https://doi.org/10.1002/path.2806

Fehlmann T, Kahraman M, Ludwig N et al (2020) Evaluating the use of circulating MicroRNA profiles for lung cancer detection in symptomatic patients. JAMA Oncol 6:714–723. https://doi.org/10.1001/jamaoncol.2020.0001

Flem-Karlsen K, Fodstad Ø, Nunes-Xavier CE (2020) B7-H3 immune checkpoint protein in human cancer. Curr Med Chem 27:4062–4086. https://doi.org/10.2174/0929867326666190517115515

Fordham JB, Naqvi AR, Nares S (2015) Regulation of miR-24, miR-30b, and miR-142-3p during macrophage and dendritic cell differentiation potentiates innate immunity. J Leukoc Biol 98:195–207. https://doi.org/10.1189/jlb.1a1014-519rr

Foss FM, Querfeld C, Kim YH et al (2018) Ph 1 study of MRG-106, an inhibitor of miR-155, in CTCL. J Clin Oncol 36:2511. https://doi.org/10.1200/JCO.2018.36.15_suppl.2511

Fu Y, Li F, Zhang P et al (2019) Myrothecine a modulates the proliferation of HCC cells and the maturation of dendritic cells through downregulating miR-221. Int Immunopharmacol 75: 105783. https://doi.org/10.1016/j.intimp.2019.105783

Gao X, Han D, Fan W (2016) Down-regulation of RBP-J mediated by microRNA-133a suppresses dendritic cells and functions as a potential tumor suppressor in osteosarcoma. Exp Cell Res 349: 264–272. https://doi.org/10.1016/j.yexcr.2016.10.019

Gao Y, Fang P, Li WJ et al (2020) LncRNA NEAT1 sponges miR-214 to regulate M2 macrophage polarization by regulation of B7-H3 in multiple myeloma. Mol Immunol 117:20–28. https://doi.org/10.1016/j.molimm.2019.10.026

Giraldo NA, Becht E, Vano Y et al (2015) The immune response in cancer: from immunology to pathology to immunotherapy. Virchows Arch 467:127–135. https://doi.org/10.1007/s00428-015-1787-7

Gong AY, Zhou R, Hu G et al (2010) Cryptosporidium parvum induces B7-H1 expression in cholangiocytes by down-regulating microRNA-513. J Infect Dis 201:160–169. https://doi.org/10.1086/648589

Gordon S, Martinez FO (2010) Alternative activation of macrophages: mechanism and functions. Immunity 32:593–604. https://doi.org/10.1016/j.immuni.2010.05.007

Grivennikov SI, Greten FR, Karin M (2010) Immunity, inflammation, and cancer. Cell 140:883–899. https://doi.org/10.1016/j.cell.2010.01.025

Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100:57–70. https://doi.org/10.1016/s0092-8674(00)81683-9

Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144:646–674. https://doi.org/10.1016/j.cell.2011.02.013

Harrison RA, Tummala S, de Groot J (2020) Neurologic toxicities of cancer immunotherapies: a review. Curr Neurol Neurosci Rep 20:27. https://doi.org/10.1007/s11910-020-01038-2

Heijnen HFG, Schiel AE, Fijnheer R et al (1999) Activated platelets release two types of membrane vesicles: microvesicles by surface shedding and exosomes derived from exocytosis of multivesicular bodies and α-granules. Blood 94:3791–3799. https://doi.org/10.1182/blood.v94.11.3791

Hirschberger S, Hinske LC, Kreth S (2018) MiRNAs: dynamic regulators of immune cell functions in inflammation and cancer. Cancer Lett 431:11–21. https://doi.org/10.1016/j.canlet.2018.05.020

Hong DS, Kang Y-K, Borad M et al (2020) Phase 1 study of MRX34, a liposomal miR-34a mimic, in patients with advanced solid tumours. Br J Cancer 122:1630–1637. https://doi.org/10.1038/s41416-020-0802-1

Hosseini A, Gharibi T, Marofi F et al (2020) CTLA-4: from mechanism to autoimmune therapy. Int Immunopharmacol 80:106221. https://doi.org/10.1016/j.intimp.2020.106221

Hsieh CH, Tai SK, Yang MH (2018) Snail-overexpressing cancer cells promote M2-like polarization of tumor-associated macrophages by delivering MiR-21-abundant exosomes. Neoplasia (United States) 20:775–788. https://doi.org/10.1016/j.neo.2018.06.004

Huang F, Wang B, Zeng J et al (2018) MicroRNA-374b inhibits liver cancer progression via down regulating programmed cell death-1 expression on cytokine-induced killer cells. Oncol Lett 15: 4797–4804. https://doi.org/10.3892/ol.2018.7951

Huang R, Chen X, Long Y, Chen R (2019) MIR-31 promotes Th22 differentiation through targeting Bach2 in coronary heart disease. Biosci Rep 39:1–11. https://doi.org/10.1042/BSR20190986

Huffaker TB, Lee SH, Tang WW et al (2017) Antitumor immunity is defective in T cell–specific microRNA-155– deficient mice and is rescued by immune checkpoint blockade. J Biol Chem 292:18530–18541. https://doi.org/10.1074/jbc.M117.808121

Ishii H, Vodnala SK, Achyut BR et al (2018) MiR-130a and miR-145 reprogram Gr-1+CD11b+ myeloid cells and inhibit tumor metastasis through improved host immunity. Nat Commun 9: 2611. https://doi.org/10.1038/s41467-018-05023-9

Ji Y, Wrzesinski C, Yu Z et al (2015) miR-155 augments CD8+ T-cell antitumor activity in lymphoreplete hosts by enhancing responsiveness to homeostatic γc cytokines. Proc Natl Acad Sci U S A 112:476–481. https://doi.org/10.1073/pnas.1422916112

Jia L, Wu C (2014) The biology and functions of Th22 cells. Adv Exp Med Biol 841:209–230. https://doi.org/10.1007/978-94-017-9487-9_8

Jia L, Xi Q, Wang H et al (2017) miR-142-5p regulates tumor cell PD-L1 expression and enhances anti-tumor immunity. Biochem Biophys Res Commun 488:425–431. https://doi.org/10.1016/j.bbrc.2017.05.074

Jiang S, Li C, Olive V et al (2011) Molecular dissection of the miR-17-92 cluster's critical dual roles in promoting Th1 responses and preventing inducible Treg differentiation. Blood 118:5487–5497. https://doi.org/10.1182/blood-2011-05-355644

Jiang J, Gao Q, Gong Y et al (2018) MiR-486 promotes proliferation and suppresses apoptosis in myeloid cells by targeting Cebpa in vitro. Cancer Med 7:4627–4638. https://doi.org/10.1002/cam4.1694

Jiang M, Zhang W, Zhang R et al (2020) Cancer exosome-derived miR-9 and miR-181a promote the development of early-stage MDSCs via interfering with SOCS3 and PIAS3 respectively in breast cancer. Oncogene 39:4681–4694. https://doi.org/10.1038/s41388-020-1322-4

Joyce JA, Fearon DT (2015) T cell exclusion, immune privilege, and the tumor microenvironment. Science (80-) 348:74–80. https://doi.org/10.1126/science.aaa6204

Katakura S, Kobayashi N, Hashimoto H et al (2020) MicroRNA-200b is a potential biomarker of the expression of PD-L1 in patients with lung cancer. Thorac Cancer 11:2975–2982. https://doi.org/10.1111/1759-7714.13653

Kawana S, Saito R, Miki Y et al (2021) Suppression of tumor immune microenvironment via microRNA-1 after epidermal growth factor receptor-tyrosine kinase inhibitor resistance acquirement in lung adenocarcinoma. Cancer Med 10:718–727. https://doi.org/10.1002/cam4.3639

Kim TD, Lee SU, Yun S et al (2011) Human microRNA-27a*targets Prf1 and GzmB expression to regulate NK-cell cytotoxicity. Blood 118:5476–5486. https://doi.org/10.1182/blood-2011-04-347526

Kim N, Kim M, Yun S et al (2014) MicroRNA-150 regulates the cytotoxicity of natural killers by targeting perforin-1. J Allergy Clin Immunol 134:195–203. https://doi.org/10.1016/j.jaci.2014.02.018

Kogure A, Kosaka N, Ochiya T (2019) Cross-talk between cancer cells and their neighbors via miRNA in extracellular vesicles: an emerging player in cancer metastasis. J Biomed Sci 26:1–8. https://doi.org/10.1186/s12929-019-0500-6

Kohlhaas S, Garden OA, Scudamore C et al (2009) Cutting edge: the Foxp3 target miR-155 contributes to the development of regulatory T cells. J Immunol 182:2578–2582. https://doi.org/10.4049/jimmunol.0803162

Kosaka A, Ohkuri T, Ikeura M et al (2015) Transgene-derived overexpression of miR-17-92 in CD8+ T-cells confers enhanced cytotoxic activity. Biochem Biophys Res Commun 458:549–554. https://doi.org/10.1016/j.bbrc.2015.02.003

Kuen DS, Kim BS, Chung Y (2020) Il-17-producing cells in tumor immunity: friends or foes? Immune Netw 20:1–20. https://doi.org/10.4110/in.2020.20.e6

Kuo G, Wu CY, Yang HY (2019) MiR-17-92 cluster and immunity. J Formos Med Assoc 118:2–6. https://doi.org/10.1016/j.jfma.2018.04.013

Kwak HJ, Kim YJ, Chun KR et al (2011) Downregulation of Spry2 by miR-21 triggers malignancy in human gliomas. Oncogene 30:2433–2442. https://doi.org/10.1038/onc.2010.620

Labeur MS, Roters B, Pers B et al (1999) Generation of tumor immunity by bone marrow-derived dendritic cells correlates with dendritic cell maturation stage. J Immunol 162:168–175

Li J, Yue L, Wang H et al (2016) Th17 cells exhibit antitumor effects in MDS possibly through augmenting functions of CD8+ T cells. J Immunol Res 2016:9404705. https://doi.org/10.1155/2016/9404705

Li Z, Suo B, Long G et al (2020) Exosomal miRNA-16-5p derived from M1 macrophages enhances T cell-dependent immune response by regulating PD-L1 in gastric cancer. Front Cell Dev Biol 8:1362. https://doi.org/10.3389/fcell.2020.572689

Liang X, Liu Y, Mei S et al (2015) MicroRNA-22 impairs anti-tumor ability of dendritic cells by targeting p38. PLoS One 10:1–15. https://doi.org/10.1371/journal.pone.0121510

Liu Q, Zhang M, Jiang X et al (2011) MiR-223 suppresses differentiation of tumor-induced CD11b +Gr1 +myeloid-derived suppressor cells from bone marrow cells. Int J Cancer 129:2662–2673. https://doi.org/10.1002/ijc.25921

Liu X, Robinson SN, Setoyama T et al (2014) FOXP3 is a direct target of miR15a/16 in umbilical cord blood regulatory T cells. Bone Marrow Transplant 49:793–799. https://doi.org/10.1038/bmt.2014.57

Liu J, Fan L, Yu H et al (2019) Endoplasmic reticulum stress causes liver cancer cells to release exosomal miR-23a-3p and up-regulate programmed death ligand 1 expression in macrophages. Hepatology 70:241–258. https://doi.org/10.1002/hep.30607

Lone SN, Bhat AA, Wani NA, et al (2021) miRNAs as novel immunoregulators in cancer. Semin Cell Dev Biol 26:S1084-9521(21)00086–0. https://doi.org/10.1016/j.semcdb.2021.04.013

Lu L-F, Boldin MP, Chaudhry A et al (2010) Function of miR-146a in controlling Treg cell-mediated regulation of Th1 responses. Cell 142:914–929. https://doi.org/10.1016/j.cell.2010.08.012

Ma S, Liu M, Xu Z et al (2016) A double feedback loop mediated by microRNA-23a/27a/24-2 regulates M1 versus M2 macrophage polarization and thus regulates cancer progression. Oncotarget 7:13502–13519. https://doi.org/10.18632/oncotarget.6284

Malekghasemi S, Majidi J, Baghbanzadeh A et al (2020) Tumor-associated macrophages: protumoral macrophages in inflammatory tumor microenvironment. Adv Pharm Bull 10:556–565. https://doi.org/10.34172/apb.2020.066

Martinez-Nunez RT, Louafi F, Sanchez-Elsner T (2011) The interleukin 13 (IL-13) pathway in human macrophages is modulated by microRNA-155 via direct targeting of interleukin 13 receptor α1 (IL13Rα1). J Biol Chem 286:1786–1794. https://doi.org/10.1074/jbc.M110.169367

Mashima R (2015) Physiological roles of miR-155. Immunology 145:323–333. https://doi.org/10.1111/imm.12468

Matlung HL, Szilagyi K, Barclay NA, van den Berg TK (2017) The CD47-SIRPα signaling axis as an innate immune checkpoint in cancer. Immunol Rev 276:145–164. https://doi.org/10.1111/imr.12527

Mittelbrunn M, Gutiérrez-Vázquez C, Villarroya-Beltri C et al (2011) Unidirectional transfer of microRNA-loaded exosomes from T cells to antigen-presenting cells. Nat Commun 2:282. https://doi.org/10.1038/ncomms1285

Mollaei H, Safaralizadeh R, Rostami Z (2019) MicroRNA replacement therapy in cancer. J Cell Physiol 234:12369–12384. https://doi.org/10.1002/jcp.28058

Monnot GC, Martinez-Usatorre A, Lanitis E et al (2020) miR-155 overexpression in OT-1 CD8+ T cells improves anti-tumor activity against low-affinity tumor antigen. Mol Ther - Oncolytics 16:111–123. https://doi.org/10.1016/j.omto.2019.12.008

Montecalvo A, Larregina AT, Shufesky WJ et al (2012) Mechanism of transfer of functional microRNAs between mouse dendritic cells via exosomes. Blood 119:756–766. https://doi.org/10.1182/blood-2011-02-338004

Morelli AE, Thomson AW (2007) Tolerogenic dendritic cells and the quest for transplant tolerance. Nat Rev Immunol 7:610–621. https://doi.org/10.1038/nri2132

Muntyanu A, Netchiporouk E, Gerstein W et al (2021) Cutaneous immune-related adverse events (irAEs) to immune checkpoint inhibitors: a dermatology perspective on management. J Cutan Med Surg 25:59–76. https://doi.org/10.1177/1203475420943260

Murata Y, Saito Y, Kotani T, Matozaki T (2018) CD47-signal regulatory protein α signaling system and its application to cancer immunotherapy. Cancer Sci 109:2349–2357. https://doi.org/10.1111/cas.13663

Nahid MA, Pauley KM, Satoh M, Chan EKL (2009) miR-146a is critical for endotoxin-induced tolerance: implication in innate immunity. J Biol Chem 284:34590–34599. https://doi.org/10.1074/jbc.M109.056317

Naqvi AR, Fordham JB, Nares S (2015) miR-24, miR-30b, and miR-142-3p regulate phagocytosis in myeloid inflammatory cells. J Immunol 194:1916–1927. https://doi.org/10.4049/jimmunol.1401893

Neelapu SS, Locke FL, Bartlett NL et al (2017) Axicabtagene Ciloleucel CAR T-cell therapy in refractory large B-cell lymphoma. N Engl J Med 377:2531–2544. https://doi.org/10.1056/nejmoa1707447

Neviani P, Wise PM, Murtadha M et al (2019) Natural killer–derived exosomal miR-186 inhibits neuroblastoma growth and immune escape mechanisms. Cancer Res 79:1151–1164. https://doi.org/10.1158/0008-5472.CAN-18-0779

Nishimura M, Jung EJ, Shah MY et al (2013) Therapeutic synergy between microRNA and siRNA in ovarian cancer treatment. Cancer Discov 3:1302–1315. https://doi.org/10.1158/2159-8290.CD-13-0159

Nygren MK, Tekle C, Ingebrigtsen VA et al (2014) Identifying microRNAs regulating B7-H3 in breast cancer: the clinical impact of microRNA-29c. Br J Cancer 110:2072–2080. https://doi.org/10.1038/bjc.2014.113

Ohno M, Ohkuri T, Kosaka A et al (2013) Expression of miR-17-92 enhances anti-tumor activity of T-cells transduced with the anti-EGFRvIII chimeric antigen receptor in mice bearing human GBM xenografts. J Immunother Cancer 1:21. https://doi.org/10.1186/2051-1426-1-21

Okada H, Kohanbash G, Lotze MT (2010) MicroRNAs in immune regulation-opportunities for cancer immunotherapy. Int J Biochem Cell Biol 42:1256–1261. https://doi.org/10.1016/j.biocel.2010.02.002

Okoye IS, Coomes SM, Pelly VS et al (2014) MicroRNA-containing T-regulatory-cell-derived exosomes suppress pathogenic T helper 1 cells. Immunity 41:89–103. https://doi.org/10.1016/j.immuni.2014.05.019

Omar HA, El-Serafi AT, Hersi F et al (2019) Immunomodulatory MicroRNAs in cancer: targeting immune checkpoints and the tumor microenvironment. FEBS J 286:3540–3557. https://doi.org/10.1111/febs.15000

Pardoll DM (2012) The blockade of immune checkpoints in cancer immunotherapy. Nat Rev Cancer 12:252–264. https://doi.org/10.1038/nrc3239

Park H, Huang X, Lu C et al (2015) MicroRNA-146a and microRNA-146b regulate human dendritic cell apoptosis and cytokine production by targeting TRAF6 and IRAK1 proteins. J Biol Chem 290:2831–2841. https://doi.org/10.1074/jbc.M114.591420

Park JE, Dutta B, Tse SW et al (2019) Hypoxia-induced tumor exosomes promote M2-like macrophage polarization of infiltrating myeloid cells and microRNA-mediated metabolic shift. Oncogene 38:5158–5173. https://doi.org/10.1038/s41388-019-0782-x

Peng X-X, Yu R, Wu X et al (2020) Correlation of plasma exosomal microRNAs with the efficacy of immunotherapy in EGFR/ALK wild-type advanced non-small cell lung cancer. J Immunother Cancer 8:e000376. https://doi.org/10.1136/jitc-2019-000376

Pesce S, Greppi M, Ferretti E et al (2020) miRNAs in NK cell-based immune responses and cancer immunotherapy. Front Cell Dev Biol 8:1–13. https://doi.org/10.3389/fcell.2020.00119

Pol J, Kroemer G, Galluzzi L (2016) First oncolytic virus approved for melanoma immunotherapy. Onco Targets Ther 5:e1115641. https://doi.org/10.1080/2162402x.2015.1115641

Presotto EM, Rastrelli G, Desideri I et al (2020) Endocrine toxicity in cancer patients treated with nivolumab or pembrolizumab: results of a large multicentre study. J Endocrinol Investig 43:337–345. https://doi.org/10.1007/s40618-019-01112-8

Qin H, Holdbrooks AT, Liu Y et al (2012) SOCS3 deficiency promotes M1 macrophage polarization and inflammation. J Immunol 189:3439–3448. https://doi.org/10.4049/jimmunol.1201168

Ranji N, Sadeghizadeh M, Shokrgozar MA et al (2013) MiR-17-92 cluster: an apoptosis inducer or proliferation enhancer. Mol Cell Biochem 380:229–238. https://doi.org/10.1007/s11010-013-1678-7

Rashdan S, Minna JD, Gerber DE (2018) Diagnosis and management of pulmonary toxicity associated with cancer immunotherapy. Lancet Respir Med 6:472–478. https://doi.org/10.1016/S2213-2600(18)30172-3

Rasoolnezhad M, Safaralizadeh R, Hosseinpourfeizi MA et al (2021) MiRNA-138–5p: a strong tumor suppressor targeting PD-L-1 inhibits proliferation and motility of breast cancer cells and induces apoptosis. Eur J Pharmacol 896:173933. https://doi.org/10.1016/j.ejphar.2021.173933

Raue R, Frank A-C, Syed SN, Brüne B (2021) Therapeutic targeting of microRNAs in the tumor microenvironment. Int J Mol Sci 22:2210. https://doi.org/10.3390/ijms22042210

Ren W, Hou J, Yang C et al (2019a) Extracellular vesicles secreted by hypoxia pre-challenged mesenchymal stem cells promote non-small cell lung cancer cell growth and mobility as well as macrophage M2 polarization via miR-21-5p delivery. J Exp Clin Cancer Res 38:1–14. https://doi.org/10.1186/s13046-019-1027-0

Ren WH, Zhang XR, Li WB et al (2019b) Exosomal miRNA-107 induces myeloid-derived suppressor cell expansion in gastric cancer. Cancer Manag Res 11:4023–4040. https://doi.org/10.2147/CMAR.S198886

Rigoutsos I, Huynh T, Miranda K et al (2006) Short blocks from the noncoding parts of the human genome have instances within nearly all known genes and relate to biological processes. Proc Natl Acad Sci U S A 103:6605–6610. https://doi.org/10.1073/pnas.0601688103

Rigoutsos I, Lee SK, Nam SY et al (2017) N-BLR, a primate-specific non-coding transcript leads to colorectal cancer invasion and migration. Genome Biol 18:1–21. https://doi.org/10.1186/s13059-017-1224-0

Rodriguez A, Vigorito E, Clare S et al (2007) Requirement of bic/microRNA-155 for normal immune function. Science (80-) 316:608–611. https://doi.org/10.1126/science.1139253

Romagnani S (2000) T-cell subsets (Th1 versus Th2). Ann Allergy Asthma Immunol 85:9–18. https://doi.org/10.1016/S1081-1206(10)62426-X

Rosenberg SA et al (2015) Use of tumor infiltrating lymphocytes for the treatment of melanoma. N Engl J Med 281:1028–1034

Rowshanravan B, Halliday N, Sansom DM (2018) CTLA-4: a moving target in immunotherapy. Blood 131:58–67. https://doi.org/10.1182/blood-2017-06-741033

Sánchez-González I, Bobien A, Molnar C et al (2020) MiR-149 suppresses breast cancer metastasis by blocking paracrine interactions with macrophages. Cancer Res 80:1330–1341. https://doi.org/10.1158/0008-5472.CAN-19-1934

Sathipati YS, Ho SY (2018) Identifying a miRNA signature for predicting the stage of breast cancer. Sci Rep 8:1–11. https://doi.org/10.1038/s41598-018-34604-3

Shinohara H, Kuranaga Y, Kumazaki M et al (2017) Regulated polarization of tumor-associated macrophages by miR-145 via colorectal cancer–derived extracellular vesicles. J Immunol 199:1505–1515. https://doi.org/10.4049/jimmunol.1700167

Sica A, Erreni M, Allavena P, Porta C (2015) Macrophage polarization in pathology. Cell Mol Life Sci 72:4111–4126. https://doi.org/10.1007/s00018-015-1995-y

Simanovich E, Brod V, Rahat MM, Rahat MA (2018) Function of miR-146a-5p in tumor cells as a regulatory switch between cell death and angiogenesis: macrophage therapy revisited. Front Immunol 8:1931. https://doi.org/10.3389/fimmu.2017.01931

Smyth LA, Boardman DA, Tung SL et al (2015) MicroRNAs affect dendritic cell function and phenotype. Immunology 144:197–205. https://doi.org/10.1111/imm.12390

Spranger S, Gajewski TF (2018) Mechanisms of tumor cell-intrinsic immune evasion. Annu Rev Cancer Biol 2:213–228. https://doi.org/10.1146/annurev-cancerbio-030617-050606

Sullivan RP, Leong JW, Fehniger TA (2013) MicroRNA regulation of natural killer cells. Front Immunol 4:1–12. https://doi.org/10.3389/fimmu.2013.00044

Swann JB, Smyth MJ (2007) Immune surveillance of tumors. J Clin Invest 117:1137–1146. https://doi.org/10.1172/JCI31405

Syed SN, Frank AC, Raue R, Brüne B (2019) MicroRNA-A tumor Trojan horse for tumor-associated macrophages. Cell 8:1–25. https://doi.org/10.3390/cells8121482

Taganov KD, Boldin MP, Chang KJ, Baltimore D (2006) NF-κB-dependent induction of microRNA miR-146, an inhibitor targeted to signaling proteins of innate immune responses. Proc Natl Acad Sci U S A 103:12481–12486. https://doi.org/10.1073/pnas.0605298103

Taghikhani A, Hassan ZM, Ebrahimi M, Moazzeni SM (2019) microRNA modified tumor-derived exosomes as novel tools for maturation of dendritic cells. J Cell Physiol 234:9417–9427. https://doi.org/10.1002/jcp.27626

Takano Y, Masuda T, Iinuma H et al (2017) Circulating exosomal microRNA-203 is associated with metastasis possibly via inducing tumor-associated macrophages in colorectal cancer. Oncotarget 8:78598–78613. https://doi.org/10.18632/oncotarget.20009

Talekar M, Trivedi M, Shah P et al (2016) Combination wt-p53 and microRNA-125b transfection in a genetically engineered lung cancer model using dual CD44/EGFR-targeting nanoparticles. Mol Ther 24:759–769. https://doi.org/10.1038/mt.2015.225

Tian J, Rui K, Tang X et al (2015) MicroRNA-9 regulates the differentiation and function of myeloid-derived suppressor cells via targeting Runx1. J Immunol 195:1301–1311. https://doi.org/10.4049/jimmunol.1500209

Tung SL, Boardman DA, Sen M et al (2018) Regulatory T cell-derived extracellular vesicles modify dendritic cell function. Sci Rep 8:1–12. https://doi.org/10.1038/s41598-018-24531-8

Turiák L, Misják P, Szabó TG et al (2011) Proteomic characterization of thymocyte-derived microvesicles and apoptotic bodies in BALB/c mice. J Proteome 74:2025–2033. https://doi.org/10.1016/j.jprot.2011.05.023

van Zandwijk N, Pavlakis N, Kao SC et al (2017) Safety and activity of microRNA-loaded minicells in patients with recurrent malignant pleural mesothelioma: a first-in-man, phase 1, open-label, dose-escalation study. Lancet Oncol 18:1386–1396. https://doi.org/10.1016/S1470-2045(17)30621-6

Verneau J, Sautés-Fridman C, Sun CM (2020) Dendritic cells in the tumor microenvironment: prognostic and theranostic impact. Semin Immunol 48:101410. https://doi.org/10.1016/j.smim.2020.101410

Wang L, Yi T, Kortylewski M et al (2009) IL-17 can promote tumor growth through an IL-6-Stat3 signaling pathway. J Exp Med 206:1457–1464. https://doi.org/10.1084/jem.20090207

Wang W, Sun J, Li F et al (2012) A frequent somatic mutation in CD274 3'-UTR leads to protein over-expression in gastric cancer by disrupting miR-570 binding. Hum Mutat 33:480–484. https://doi.org/10.1002/humu.22014

Wang J, Zheng S, Xin N et al (2013a) Identification of novel microRNA signatures linked to experimental autoimmune myasthenia gravis pathogenesis: down-regulated mir-145 promotes pathogenetic th17 cell response. J Neuroimmune Pharmacol 8:1287–1302. https://doi.org/10.1007/s11481-013-9498-9

Wang Z, Han J, Cui Y et al (2013b) MiRNA-21 inhibition enhances RANTES and IP-10 release in MCF-7 via PIAS3 and STAT3 signalling and causes increased lymphocyte migration. Biochem Biophys Res Commun 439:384–389. https://doi.org/10.1016/j.bbrc.2013.08.072

Wang J, Yu F, Jia X et al (2015a) MicroRNA-155 deficiency enhances the recruitment and functions of myeloid-derived suppressor cells in tumor microenvironment and promotes solid tumor growth. Int J Cancer 136:E602–E613. https://doi.org/10.1002/ijc.29151

Wang Z, Brandt S, Medeiros A et al (2015b) MicroRNA 21 is a homeostatic regulator of macrophage polarization and prevents prostaglandin e2 -mediated M2 generation. PLoS One 10:1–13. https://doi.org/10.1371/journal.pone.0115855

Wang J, Ye H, Zhang D et al (2017) Cancer-derived circulating microRNAs promote tumor angiogenesis by entering dendritic cells to degrade highly complementary microRNAs. Theranostics 7:1407–1421. https://doi.org/10.7150/thno.18262

Wang D, Tang M, Zong P et al (2018a) MiRNA-155 regulates the Th17/Treg ratio by targeting SOCS1 in severe acute pancreatitis. Front Physiol 9:1–10. https://doi.org/10.3389/fphys.2018.00686

Wang H, Zhang Y, Wu X et al (2018b) Regulation of human natural killer cell IFN-γ production by MicroRNA-146a via targeting the NF-κB signaling pathway. Front Immunol 9:293. https://doi.org/10.3389/fimmu.2018.00293

Wang T, Wang D, Yu H et al (2018c) A cancer vaccine-mediated postoperative immunotherapy for recurrent and metastatic tumors. Nat Commun 9:1–12. https://doi.org/10.1038/s41467-018-03915-4

Wang W, Liu Y, Guo J et al (2018d) miR-100 maintains phenotype of tumor-associated macrophages by targeting mTOR to promote tumor metastasis via Stat5a/IL-1ra pathway in mouse breast cancer. Oncogenesis 7:97. https://doi.org/10.1038/s41389-018-0106-y

Wang L, Hu Y-Y, Zhao J-L et al (2020a) Targeted delivery of miR-99b reprograms tumor-associated macrophage phenotype leading to tumor regression. J Immunother Cancer 8: e000517. https://doi.org/10.1136/jitc-2019-000517

Wang Y, Jia A, Bi Y et al (2020b) Targeting myeloid-derived suppressor cells in cancer immunotherapy. Cancers (Basel) 12:1–24. https://doi.org/10.3390/cancers12092626

Wei J, Nduom EK, Kong L-Y et al (2016) MiR-138 exerts anti-glioma efficacy by targeting immune checkpoints. Neuro-Oncology 18:639–648. https://doi.org/10.1093/neuonc/nov292

Wei SC, Duffy CR, Allison JP (2018) Fundamental mechanisms of immune checkpoint blockade therapy. Cancer Discov 8:1069–1086. https://doi.org/10.1158/2159-8290.CD-18-0367

Weiskopf K (2017) Cancer immunotherapy targeting the CD47/SIRPα axis. Eur J Cancer 76:100–109. https://doi.org/10.1016/j.ejca.2017.02.013

Wing K, Onishi Y, Prieto-Martin P et al (2008) CTLA-4 control over Foxp3+ regulatory T cell function. Science (80-) 322:271–275. https://doi.org/10.1126/science.1160062

Xi Q, Chen Y, Yang G-Z et al (2020) miR-128 regulates tumor cell CD47 expression and promotes anti-tumor immunity in pancreatic cancer. Front Immunol 11:890. https://doi.org/10.3389/fimmu.2020.00890

Xing Y, Ruan G, Ni H et al (2021) Tumor immune microenvironment and its related miRNAs in tumor progression. Front Immunol 12:1–16. https://doi.org/10.3389/fimmu.2021.624725

Xu S, Tao Z, Hai B et al (2016) miR-424(322) reverses chemoresistance via T-cell immune response activation by blocking the PD-L1 immune checkpoint. Nat Commun 7:11406. https://doi.org/10.1038/ncomms11406

Xu D, Han Q, Hou Z et al (2017a) MIR-146a negatively regulates NK cell functions via STAT1 signaling. Cell Mol Immunol 14:712–720. https://doi.org/10.1038/cmi.2015.113

Xu Z, Ji J, Xu J et al (2017b) MiR-30a increases MDSC differentiation and immunosuppressive function by targeting SOCS3 in mice with B-cell lymphoma. FEBS J 284:2410–2424. https://doi.org/10.1111/febs.14133

Yamashita T, Takahashi Y, Takakura Y (2018) Possibility of exosome-based therapeutics and challenges in production of exosomes eligible for therapeutic application. Biol Pharm Bull 41: 835–842. https://doi.org/10.1248/bpb.b18-00133

Yang P, Li QJ, Feng Y et al (2012) TGF-β-miR-34a-CCL22 signaling-induced treg cell recruitment promotes venous metastases of HBV-positive hepatocellular carcinoma. Cancer Cell 22:291–303. https://doi.org/10.1016/j.ccr.2012.07.023

Yang Q, Cao W, Wang Z et al (2018) Regulation of cancer immune escape: the roles of miRNAs in immune checkpoint proteins. Cancer Lett 431:73–84. https://doi.org/10.1016/j.canlet.2018.05.015

Yang Q, Li J, Hu Y et al (2019) MiR-218-5p suppresses the killing effect of natural killer cell to lung adenocarcinoma by targeting SHMT1. Yonsei Med J 60:500–508. https://doi.org/10.3349/ymj.2019.60.6.500

Yang S, Wei W, Zhao Q (2020) B7-H3, a checkpoint molecule, as a target for cancer immunotherapy. Int J Biol Sci 16:1767–1773. https://doi.org/10.7150/ijbs.41105

Yao X, Tu Y, Xu Y et al (2020) Endoplasmic reticulum stress-induced exosomal miR-27a-3p promotes immune escape in breast cancer via regulating PD-L1 expression in macrophages. J Cell Mol Med 24:9560–9573. https://doi.org/10.1111/jcmm.15367

Yatim N, Cullen S, Albert ML (2017) Dying cells actively regulate adaptive immune responses. Nat Rev Immunol 17:262–275. https://doi.org/10.1038/nri.2017.9

Ye S-B, Zhang H, Cai T-T et al (2016) Exosomal miR-24-3p impedes T-cell function by targeting FGF11 and serves as a potential prognostic biomarker for nasopharyngeal carcinoma. J Pathol 240:329–340. https://doi.org/10.1002/path.4781

Yee D, Shah KM, Coles MC et al (2017) MicroRNA-155 induction via TNF-α and IFN-γ suppresses expression of programmed death ligand-1 (PD-L1) in human primary cells. J Biol Chem 292:20683–20693. https://doi.org/10.1074/jbc.M117.809053

Yin Z, Deng Z, Zhao W, Cao Z (2018) Searching synergistic dose combinations for anticancer drugs. Front Pharmacol 9:535. https://doi.org/10.3389/fphar.2018.00535

Yoon CJ, Chang MS, Kim DH, et al (2020) Epstein–Barr virus-encoded miR-BART5-5p upregulates PD-L1 through PIAS3/pSTAT3 modulation, worsening clinical outcomes of PD-L1-positive gastric carcinomas. Gastric Cancer 23:780–795. https://doi.org/10.1007/s10120-020-01059-3

Yu J, Freud AG, Caligiuri MA (2013) Location and cellular stages of natural killer cell development. Trends Immunol 34:573–582. https://doi.org/10.1016/j.it.2013.07.005

Zhang M, Liu F, Jia H et al (2011) Inhibition of MicroRNA let-7i depresses maturation and functional state of dendritic cells in response to lipopolysaccharide stimulation via targeting suppressor of cytokine signaling 1. J Immunol 187:1674–1683. https://doi.org/10.4049/jimmunol.1001937

Zhang Y, Zhang M, Zhong M et al (2013) Expression profiles of miRNAs in polarized macrophages. Int J Mol Med 31:797–802. https://doi.org/10.3892/ijmm.2013.1260

Zhang A, Qian Y, Ye Z et al (2017) Cancer-associated fibroblasts promote M2 polarization of macrophages in pancreatic ductal adenocarcinoma. Cancer Med 6:463–470. https://doi.org/10.1002/cam4.993

Zhang LL, Zhang LF, Shi YB (2018a) miR-24 inhibited the killing effect of natural killer cells to colorectal cancer cells by downregulating Paxillin. Biomed Pharmacother 101:257–263. https://doi.org/10.1016/j.biopha.2018.02.024

Zhang T, Zhang Z, Li F et al (2018b) miR-143 regulates memory T cell differentiation by reprogramming T cell metabolism. J Immunol 201:2165–2175. https://doi.org/10.4049/jimmunol.1800230

Zhang Z, Xue Z, Liu Y et al (2018c) MicroRNA-181c promotes Th17 cell differentiation and mediates experimental autoimmune encephalomyelitis. Brain Behav Immun 70:305–314. https://doi.org/10.1016/j.bbi.2018.03.011

Zhang M, Gao D, Shi Y et al (2019) miR-149-3p reverses CD8+ T-cell exhaustion by reducing inhibitory receptors and promoting cytokine secretion in breast cancer cells. Open Biol 9:190061. https://doi.org/10.1098/rsob.190061

Zhang M, Shi Y, Zhang Y et al (2020) miRNA-5119 regulates immune checkpoints in dendritic cells to enhance breast cancer immunotherapy. Cancer Immunol Immunother 69:951–967. https://doi.org/10.1007/s00262-020-02507-w

Zhao L, Yu H, Yi S et al (2016) The tumor suppressor miR-138-5p targets PD-L1 in colorectal cancer. Oncotarget 7:45370–45384. https://doi.org/10.18632/oncotarget.9659

Zhou M, Chen J, Zhou L et al (2014) Pancreatic cancer derived exosomes regulate the expression of TLR4 in dendritic cells via miR-203. Cell Immunol 292:65–69. https://doi.org/10.1016/j.cellimm.2014.09.004

Zhu SY, Wu QY, Zhang CX et al (2018) miR-20a inhibits the killing effect of natural killer cells to cervical cancer cells by downregulating RUNX1. Biochem Biophys Res Commun 505:309–316. https://doi.org/10.1016/j.bbrc.2018.09.102

Chapter 4
Machine Learning Based Methods and Best Practices of microRNA-Target Prediction and Validation

Neetika Nath and Stefan Simm

Abstract Within the last years, more and more noncoding RNAs (ncRNAs) became the focal point to understand cell regulatory mechanisms because one class of ncRNAs, microRNAs (miRNAs), plays an essential role in translation repression or degradation of specific mRNAs and is implicated in disease etiology. miRNAs can serve as oncomiRs (oncogenic miRNAs) and tumor suppressor miRNAs, thus, miRNA therapeutics in clinical trials have become a vital component with respect to cancer treatment. To circumvent side-effects and allow an accurate effect it is crucial to accurately predict miRNAs and their mRNA targets. Over the last two decades, different approaches for miRNA prediction as well as miRNA target prediction have been developed and improved based on sequence and structure features. Nowadays, the abundance of high-throughput sequencing data and databases of miRNAs and miRNA targets from different species allow the training, testing, and validation of predicted miRNAs and miRNA targets with machine learning methods. This book chapter focuses on the important requirements for miRNA target prediction tools using ML like common features used for miRNA-binding site prediction. Furthermore, best practices for the prediction and validation of miRNA–mRNA targets are presented and set in the context of possible applications for cancer diagnosis and therapeutics.

Keywords Machine learning · Best practices · Cancer · miRNA–mRNA target prediction · miRNA validation

4.1 Introduction of miRNAs and Their Role in Cancer

The progressions of the post-genomic era have paved the way to focus on regulation based on RNomics including protein-coding messenger RNAs (mRNAs) and non-coding RNAs (ncRNAs) (Huttenhofer et al. 2002). Among these ncRNAs,

N. Nath · S. Simm (✉)
Institute of Bioinformatics, University Medicine Greifswald, Greifswald, Germany
e-mail: neetika.nath@uni-greifswald.de; stefan.simm@uni-greifswald.de

U. Schmitz et al. (eds.), *Systems Biology of MicroRNAs in Cancer*, Advances in Experimental Medicine and Biology 1385,
https://doi.org/10.1007/978-3-031-08356-3_4

microRNAs (miRNAs; 19–25 nucleotides in length) are receiving increasing attention as they regulate gene expression at the posttranscriptional level (Chakraborty et al. 2018). In principle, miRNAs are transcribed through RNA polymerase II, spliced, capped, and polyadenylated, resulting in the primary miRNA (pri-miRNA). Subsequently, pri-miRNAs are processed via the endonuclease Drosha leading to a specific hairpin structure referred to as "precursor" miRNAs (Adams 2017). Once processed from the hairpin, the mature miRNA is loaded in the Argonaute protein of the RNA-induced silencing complex (RISC) and pairs with mRNAs to direct posttranscriptional repression. To understand the functional regulation of miRNAs, it is vital to identify the target mRNAs (Bartel 2009). Predictions of potential miRNA–mRNA targets are based on specific properties of miRNA response elements (MREs) and evolutionary conservation of target sites (Simkin et al. 2020). As it is noted that miRNAs genes are often located in specific key sites like cancer-associated regions (Chakraborty et al. 2018) the functional understanding has the potential to be used in cancer treatment applications (Box 4.1).

Box 4.1 Types of miRNAs in Cancer

Studies have conclusively demonstrated that miRNAs are deeply involved in tumor onset and progression either behaving as tumor-promoting miRNAs (oncomiRs and metastamiRs) or as tumor suppressor miRNAs (Volinia et al. 2006). For example, miR-25 acting as oncomiR in osteosarcoma is negatively regulating p27 protein expression (Wang et al. 2014). Additionally, metastamiRs are a subclass of oncomiRs associated with the acquisition of metastatic phenotypes by a metastasis-promoting or tumor suppressor inhibitory activity (White et al. 2011), for example, miR-1908 decreases the expression of the tumor suppressor PTEN in glioblastoma cells resulting in an increase in proliferation, migration, and invasion (Xia et al. 2015). Furthermore, miRNAs can function as both oncomiR and metastamiR like miR-96 regulating the TGF-β/mTOR signaling, promoting bone metastasis, and contributing to a reduced survival rate in prostate cancer (Siu et al. 2015). In contrast to oncomiRs and metastamiRs, tumor suppressor miRNAs exist, which are often downregulated in tumors and exhibit onco-suppressor properties by targeting oncoprotein coding mRNAs. For example, miR-340 is responsible for the downregulation of the posttranscriptional regulators PUM1, PUM2, and SKP2 and this is involved in the upregulation of p27 and has the opposite function of miR-25 (Fernandez et al. 2015).

Alterations in miRNA expression have been demonstrated to be associated with cancer, where guided alterations of specific miRNAs have been suggested as novel therapeutic approaches (Garzon et al. 2010). A review by Chakraborty et al. (2017) pinpoint miRNA controlling cancer stem cells and their role in carcinogenesis (Fig. 4.1). Within the last decades, cancer-related miRNAs are increasingly identified and characterized of which the majority are located in cancer-associated

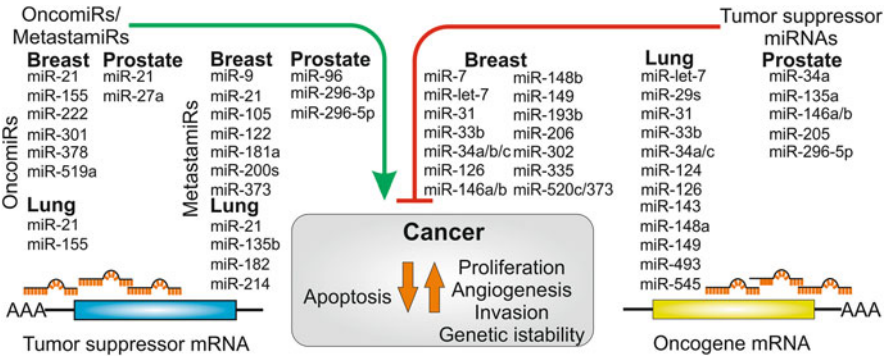


Fig. 4.1 Cancer-related miRNAs (breast, lung, and prostate) in *Homo sapiens*. The scheme represents the known tumor suppressor miRNAs (right site) inhibiting oncogene mRNA targets and oncomiRs and metastamiRs (left side) inhibiting tumor suppressor mRNAs. miRNAs were extracted from the review of Gambari et al. (2016)

genomic regions in solid tumors and hematological malignancies (Gambari et al. 2016). Further, it was confirmed that cancer-specific miRNAs can also be present in body fluids. These miRNAs play an essential role in the crosstalk between cancer cells and surrounding normal cells (Weber et al. 2010).

Besides the identification of miRNAs and their expression profiles within the cell, the prediction of target mRNAs is crucial to understanding the regulatory mechanism of the miRNAs in the context of diseases. Challenging for the prediction of miRNA–mRNA target sites is the fact that several types of miRNA binding sites exist differing in the position and localization of Watson–Crick pairings and mismatches (Bartel 2009). Additionally, the rules for miRNA–mRNA target sites vary within eukaryotic clades. In plants, most miRNA–mRNA target sites are located within the open reading frames (ORFs) of target genes and have nearly full complementarity to the miRNA (Voinnet 2009), whereas in animals, miRNA–mRNA target prediction is more challenging because of diverse binding rules (Witkos et al. 2011). First, only a few miRNA–mRNA targets show a strict complementarity in the target site and this increases the interaction complexity. Second, there is only limited knowledge about the rules as miRNA binding sites are predominantly found in 3′ untranslated regions (UTRs) of target genes and only sporadically in 5′ UTRs or ORFs (Lytle et al. 2007). Therefore within the last decades, new findings of miRNA biology and their targets including novel miRNA–mRNA target features implemented in prediction algorithms helped to increase accuracy in the prediction of miRNA–mRNA interactions (Witkos et al. 2011). Furthermore, numerous target prediction algorithms exploiting optimization function of machine learning (ML) for refining the miRNA–mRNA target predictions were developed.

4.2 miRNA Target Prediction Tools

Besides the prediction of miRNAs (Box 4.2) another big challenge is the identification of miRNA–mRNA target sites driven by sequence complementarity. Given that a genome-wide detection of new miRNAs and their target sites would be very time consuming and expensive in the wet lab, computational approaches are used to screen the genome for potential targets and their target sites. Subsequently, the high-confidence candidates can be validated by wet lab experiments. To support the user in choosing between the variety of miRNA target prediction tools two interactive guide tools are available as web services to give an overview of available miRNA target prediction tools: MT-guide (Kern et al. 2020) and Tools4miRs (Lukasik et al. 2016). These tools provide manually curated catalogs of different methods that are available for miRNA–mRNA target prediction. There are multistep questionnaires and check boxes to set distinguishing features like prediction method, target organism, target region, and prediction features.

4.2.1 Common Features for miRNA Target Prediction

The common principle of computational methods lies in the identification of complementary sequence matches between miRNAs and the identified target genes (Min and Yoon 2010). Besides the fact that perfect complementary matches between miRNA and mRNA are rarely observed, especially in mammalian organisms, the algorithms need to consider near-perfect complementary matches and further features. The commonly used features in such miRNA target prediction tools are local seed sequence comparisons, structural sequence information, and global sequence information like conservation and abundance of target sites (Peterson et al. 2014). However, these features provide limited knowledge to cover all existing special cases of the complex miRNA–mRNA target binding mechanisms in vivo for example miRNA/mRNA pairing lacking seed complementary (Chipman and Pasquinelli 2019). Additionally, the prediction of putative miRNA targets is not directly connected to biological significance, as multiple mRNA targets exist for one miRNA and (anti-)correlated expression levels have to be set into biological context (Misiewicz-Krzeminska et al. 2019).

Box 4.2 Prediction of miRNAs Using ML
Genome-wide identification of miRNAs within a species is performed using high-throughput sequencing like miRNAseq. The sequencing reads are mapped on the genome to identify known miRNAs. Expressed unannotated regions are considered for novel miRNA prediction. For the prediction via ML techniques, the general workflow is to train the classifier model using a

(continued)

Box 4.2 (continued)
positive and negative dataset for miRNAs, at which the structure of miRNAs has to be predicted beforehand using tools like Vienna RNAfold (Gruber et al. 2008). The positive set contains hairpin sequences of the miRNAs from experimentally verified databases like miRBase (Kozomara et al. 2019) or RFAM (Griffiths-Jones et al. 2003), which can be filtered based on specific features and should not contain redundancy from miRNA duplexes. Whereas, the negative set is created in silico using random segments from pre-miRNA hairpins to create pseudo miRNA duplexes. Afterward, the classifier is trained on both sets using different features like structure (stem and loop), sequence, or thermodynamic energies and can be extended using only the mature miRNA sequences as additional features. The best-trained model will be chosen based on the cross-validation results. As possible classifiers mainly support vector machines (SVM; e.g., miR-abela (Sebastian and Aggrey 2008) or Triplet-SVM (Xue et al. 2005)), Hidden Markov models (HMM; e.g., ProMirII (Nam et al. 2006), HHMMiR (Kadri et al. 2009)) or naïve Bayes methods (e.g., BayesMirFind (Yousef et al. 2006) or miR-KDE (Chang et al. 2008)) are used. Besides, also random forest (RF)-based methods like MiPred (Jiang et al. 2007) and neural networks (NN) (Zheng et al. 2020) have been implemented to predict pre-miRNAs or novel miRNAs.

Until now miRNA target prediction (Fig. 4.2) is an important, dynamic research field. Prediction tools are classified into six categories (Chu et al. 2020) based on six main features (Grimson et al. 2007; Peterson et al. 2014; Riolo et al. 2020): (1) Seed matching is the most common principle that is incorporated in many target prediction algorithms. To calculate the seed matching score, the general rule is to determine the complementarity of the miRNA seed to the mRNA. The miRNA seed sequence is 2–8 nucleotides long starting at the 5′ end and counting toward the 3′

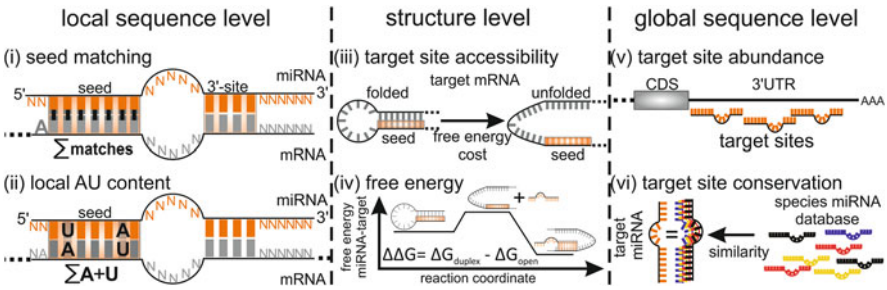


Fig. 4.2 Features for miRNA–mRNA target prediction. The scheme represents the main features for miRNA–mRNA target prediction categorized in local sequence level: seed matching (1); local AU content (2); in structure level: target site accessibility (3); free energy (4); global sequence level: target site abundance (5); and target site conservation (6). The miRNAs are visualized in orange and the mRNA target in gray

end. This region binds to the 3′ UTR of mRNAs and the efficacy of target repression is dependent on the number of consecutive base pairs (1mer, 2mer, . . .) and their complementary (Lewis et al. 2003). So far the majority of known miRNA–mRNA target sites can be classified into three types of canonical sites based on the seed matching feature (Benway and Iacomini 2018): the 7mer1A that has an adenine in position 1 at the 5′ end of miRNA, the 8mer having matched adenine in position 1 and an additional match in position 8 and the 7mer-m8 that has a match in position 8. Besides these main types, shorter seed matching regions like the 6mer, and 3′-supplementary/-compensatory sites have been observed. Thereby the 3′ part of the miRNA (positions 13–16) corresponds to nucleotides of the transcript and can compensate for a mismatch in the seed region (Grimson et al. 2007).

Within the seed matching, local abundance of adenine (A) and uracil (U) in the seed region are considered as a specific feature (Ghoshal et al. 2015). This local AU content (2) feature has intrinsically a high correlation to the seed matching feature but is less commonly used in the target prediction. (3) The feature of target site accessibility is still focusing on the seed region but provides in contrast to the local AU content information about the structure of the mRNA target site (Marin and Vanicek 2011). The accessibility is calculated using the cost of free energy to unfold the mRNA at the seed region to bind the miRNA (Kertesz et al. 2007). The second structural feature is considering (4) the free energy not only used in the unfolding likelihood of the mRNA but also as a general feature to define the stability of the miRNA–mRNA complex. Greater stability of miRNA/mRNA duplexes is reflected by lower free energy, which is influenced by other factors like length of mRNA, formation of secondary structures at thermodynamic equilibrium, binding site accessibility, and thermodynamic stability of miRNA–mRNA duplexes (Rojo Arias and Busskamp 2019). Besides these commonly used features for miRNA target site prediction focusing on the local specific seed sequence or the structural information also two features are based on a more widespread view. Here, (5) target site abundance across the whole 3′ UTR of the mRNA as well as (6) the conservation of miRNA binding sites across species are considered. The target site abundance is measured by how many target sites occur in the putative mRNA 3′UTR region and should be considered as a secondary feature to the seed matching or free energy feature (Garcia et al. 2011). In contrast, the conservation feature is based on the hypothesis that a miRNA target site, which is conserved across species, is the result of positive natural selection and this adds a functional meaning to the prediction. The method is based on the analysis of the UTR as well as the miRNA and includes different "seed types" across different clades for their calculations (Simkin et al. 2020). The inclusion of next generation sequencing (NGS) data leads to additional information like co-expression, which is used for in silico prediction (Alexiou et al. 2009).

Currently, various tools for miRNA–mRNA target prediction are available that are roughly categorized into feature driven (derived from characteristics of the mRNA sequence and/or miRNA–mRNA interaction) and statistics driven (statistical inference based on ML) (Riolo et al. 2020). The feature-driven algorithms use different features of the miRNA-target complex to identify de novo interactions,

whereas in the case of statistics-driven methods, the idea is to identify miRNA targets referenced by biological significant miRNA–mRNA duplexes from sample datasets to use the acquired information on unknown data (Witkos et al. 2011).

4.2.2 *Machine Learning Based Algorithms for miRNA Target Prediction*

The generic term machine learning (ML) comprises algorithms like logistic/lasso regression, support vector machine (SVM), random forest (RF), and neural network (NN). These algorithms can be supervised or unsupervised, at which in the case of miRNA target prediction supervised methods are mainly used. In principle, the aim of ML techniques is to train and reduce the internal error on known input datasets with labels using an optimizer function. The trained ML model can be afterward applied to new datasets to predict the class labels as output (Sidey-Gibbons and Sidey-Gibbons 2019). In contrast to the miRNA prediction (Box 4.2), miRNA–mRNA target prediction methods based on ML are trained to distinguish between target and non-target. Therefore the general outline is to identify for a set of validated targets (positive) and non-targets (negative) the putative binding sites of the miRNA, extract the feature information (free energy, seed matching, etc.), train the ML classifier and use the trained ML to decide for a separate set of miRNA–mRNA pairs between target and non-target (Parveen et al. 2019). To focus on cancer-specific miRNA–mRNA target prediction based on ML methods the search space was limited to prediction tools included in MT-guide and Tools4miRs containing *Homo sapiens* as target organism and ML as prediction method published in the time interval between 2006 and 2020 (Fig. 4.3).

Linear ML approaches, like lasso or logistic regression (miRNALasso (Wang et al. 2015), TargetThermo (Lekprasert et al. 2011), MirAncesTar (Leclercq et al. 2017), and STarMir (Rennie et al. 2014) outperform static rule-based miRNA target prediction algorithms due to their ability to dynamically adjust linear relationships in their models based on newly detected targets. The linear ML models define a prediction outcome (target vs. non-target) based on different sequences and structural features. Where tools like TargetThermo only consider seed matching and energy-based features for miRNA–mRNA target prediction, miRNALasso additionally includes co-expression data of mRNA and miRNA to predict the regulatory effect. In addition to sequence features, some tools and web servers like STarMir include experimental data like crosslinking immunoprecipitation in humans for predicting targets (V-CLIP) (Kishore et al. 2011). The tool MirAncesTar is considering the human miRNAs and their mammalian orthologues from the database miRBase (Kozomara et al. 2019) to boost the prediction accuracy of their model. In contrast to nonlinear ML techniques like SVM, RF or NN are capable of training on all available features and modeling nonlinear relationships like miRNA

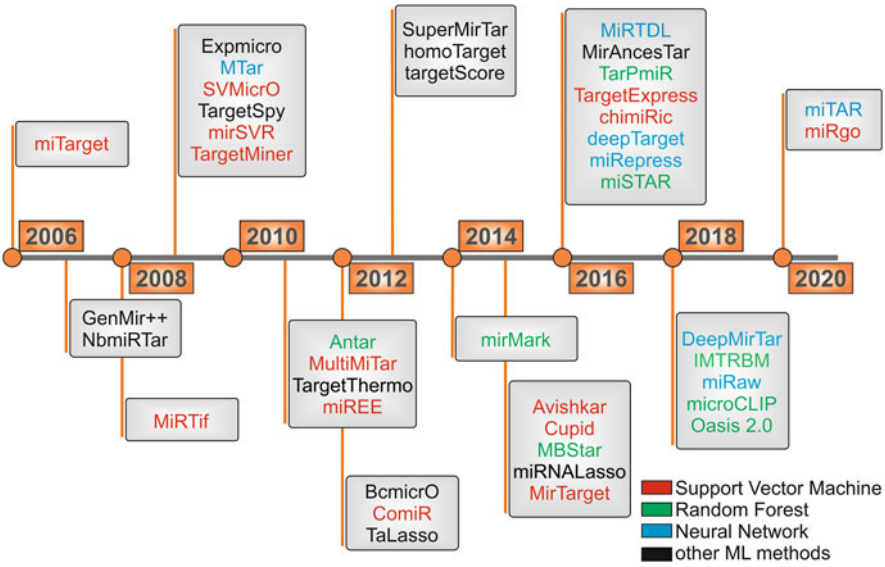


Fig. 4.3 Timeline for ML-based miRNA target prediction tools. The timeline represents the release of the first version of developed tools from 2006 to 2020 categorized in support vector machine (red), random forest (green), neural network (blue), and other methods (Bayesian inference, regression; black). Tools are limited to *Homo sapiens* including miRNA target prediction gathered from MT-guide and Tools4miRs

expression and target repression (Pelaez and Carthew 2012), which can lead to more insights into the target interactions (Schafer and Ciaudo 2020).

In the beginning, starting from 2006, most miRNA target prediction tools were based on linear ML methods including Gaussian mixture models or some nonlinear Bayesian mixture models (Expmicro (Liu et al. 2010b), GenMir++ (Huang et al. 2007), ElMMo3 (Gaidatzis et al. 2007), BcmicrO (Yue et al. 2012), TargetSpy (Sturm et al. 2010), DIANA-microT-CDS (Paraskevopoulou et al. 2013), SuperMirTar (Liu et al. 2013), TaLasso (Muniategui et al. 2012), NbmiRTar (Yousef et al. 2007), homoTarget (Ahmadi et al. 2013)). In the first 10 years of miRNA target prediction, many tools implemented SVMs beside the mixture models like Avishkar, chimiRic, ComiR, Cupid, miREE, miRgo, mirSVR, MiRTif, miTarget, MultiMiTar, SVMicrO, TargetExpress, and TargetMiner (Table 4.1). SVM-based prediction methods are still implemented today in tools like miRgo (Chu et al. 2020). In principle, SVMs use the kernel trick to map the input dataset in a higher-dimensional space through a chosen kernel function. This allows the SVM method to construct an optimized hyperplane in the feature space to maximize segregation of miRNA–mRNA targets from non-targeted interactions (Vapnik 1999).

Between 2010 and 2018 ML tools based on RF like Antar, IMTRBM, MBStar, microCLIP, mirMark, miSTAR, Oasis 2.0, TargetScan, and TarPmiR have been released (Table 4.2). The most important property of RF is to select features and

Table 4.1 SVM-based methods for miRNA target prediction ([a]...NGS input)

Tool	References	Web	Target region	Prediction features					
				1	2	3	4	5	6
miTarget	Kim et al. (2006)	W	3′ UTR	Yes	No	No	Yes	No	No
MiRTif	Yang et al. (2008)	W	Any	Yes	Yes	No	No	No	No
TargetMiner[a]	Bandyopadhyay and Mitra (2009)	Both	3′ UTR	Yes	No	No	No	No	No
SVMicrO	Liu et al. (2010a)	D	3′ UTR	Yes	No	Yes	Yes	Yes	Yes
mirSVR	Betel et al. (2010)		3′ UTR	Yes	Yes	No	No	No	No
MultiMiTar	Mitra and Bandyopadhyay (2011)	Both	3′ UTR	Yes	Yes	No	No	No	No
miREE	Reyes-Herrera et al. (2011)	W	3′ UTR	Yes	No	Yes	Yes	No	No
ComiR[a]	Zhao and Xue (2019)	Both	3′ UTR	No	No	No	No	No	No
Avishkar	Ghoshal et al. (2016)	D	Any	Yes	Yes	Yes	No	No	No
Cupid	Chiu et al. (2015)	D	3′ UTR	Yes	Yes	Yes	Yes	Yes	Yes
MirTarget	Liu and Wang (2019)	W	Any	Yes	Yes	No	Yes	Yes	No
chimiRic	Lu and Leslie (2016)	D	3′ UTR	Yes	Yes	No	No	No	No
TargetExpress[a]	Ovando-Vazquez et al. (2016)	Both	3′ UTR	No	No	No	No	No	No
miRgo	Chu et al. (2020)	W	Any	Yes	No	No	Yes	No	No

Availability (web server: W; download: D; both), target region (3′ UTR; 5′ UTR; CDS; Any), and prediction features (1: seed matching; 2: local AU content; 3: target site accessibility; 4: free energy; 5: target site abundance; 6: target site conservation) are listed

Table 4.2 RF methods for miRNA target prediction ([a]...NGS input)

Tool	References	Web	Target region	Prediction features					
				1	2	3	4	5	6
Antar	Wen et al. (2011)		3′ UTR	Yes	No	Yes	Yes	No	No
mirMark	Menor et al. (2014)	D	3′ UTR	Yes	Yes	Yes	Yes	No	Yes
MBStar	Bandyopadhyay et al. (2015)	Both	3′ UTR	Yes	No	Yes	Yes	No	No
TarPmiR	Ding et al. (2016)	D	Any	Yes	Yes	Yes	Yes	No	No
miSTAR	Van Peer et al. (2017)	W	3′ UTR	Yes	Yes	No	Yes	Yes	No
Oasis 2.0[a]	Capece et al. (2015)	W	Any	No	No	No	No	No	No
microCLIP	Paraskevopoulou et al. (2018)	D	3′ UTR, CDS	Yes	No	No	Yes	No	Yes
IMTRBM	Liu et al. (2019)	D	Any	Yes	No	Yes	Yes	No	No
TargetScan	Lewis et al. (2003)	Both	3′ UTR	Yes	No	Yes	No	No	Yes

Availability (web server: W; download: D; both), target region (3′ UTR; 5′ UTR; CDS; Any), and prediction features (1: seed matching; 2: local AU content; 3: target site accessibility; 4: free energy; 5: target site abundance; 6: target site conservation) are listed

Table 4.3 NN tools for miRNA–mRNA target prediction

Tool	References	Web	Target region	Prediction features					
				1	2	3	4	5	6
MTar	Chandra et al. (2010)		Any	Yes	Yes	No	Yes	No	No
miRepress	Ghosal et al. (2016)	Both	Any	Yes	No	Yes	No	Yes	Yes
MiRTDL	Cheng et al. (2016)	Both	Any	Yes	No	Yes	No	No	Yes
deepTarget	Lee et al. (2016)	D	Any	No	No	Yes	No	Yes	Yes
DeepMirTar	Wen et al. (2018)	D	3′ UTR	Yes	No	Yes	Yes	No	Yes
miRaw	Pla et al. (2018)	D	3′ UTR	Yes	No	Yes	Yes	No	No
miTAR	Gu et al. (2021)	D	Any	Yes	No	No	No	No	No

Availability (web server: W; download: D; Both), target region (3′ UTR; 5′ UTR; CDS; Any), and prediction features (1: seed matching; 2: local AU content; 3: target site accessibility; 4: free energy; 5: target site abundance; 6: target site conservation) are shown

notice their relevance during the model training. In principle, RF is an ensemble approach operating by building a multitude of decision trees. Several decision trees are trained with random bootstrap samples from two-thirds of the original dataset and afterward combined into a single prediction by means of voting (Svetnik et al. 2003).

Since 2016, tools using NN approaches like MTar, miRepress, MiRTDL, deepTarget, DeepMirTar, miRaw, and miTAR have been increasingly released and becoming the state-of-the-art technique in this research area (Table 4.3). Many of these NN methods for miRNA target prediction belong in the category of deep learning because they are based on multiple layers that progressively extract features from the input. Such neural nets can be subdivided into categories like convolutional neural networks (CNN), auto-encoder, or recurrent neural networks (RNN). Basically, NNs are trained iteratively over the single layers until a pattern can be identified to distinguish between target and non-target. After the input layer several hidden layers of neurons, each responsible for a linear transformation followed by a nonlinear activation function are connected and end in an output layer containing the categories (Rumelhart et al. 1986).

Because ML-based prediction tools are trained on labelled datasets, these algorithms must be trained on a positive dataset of experimentally validated miRNA–mRNA interactions and a negative dataset containing artificially generated or experimentally proven non-targets. The aim is to identify specific patterns that discriminate between miRNA–mRNA target and non-target (Sidey-Gibbons and Sidey-Gibbons 2019). Supervised ML can exclusively learn from the provided examples but is able to generalize the given information within the features and transfer them to unknown results for classification (Kim et al. 2006). The fundamental importance of the training is to provide good quality datasets covering special cases and general rules (Selbach et al. 2008). The positive dataset is extracted from publicly available databases like miRBase (Kozomara et al. 2019), MiRTarBase (Hsu et al. 2011), DIANA-TarBase (Karagkouni et al. 2018), and miRecords (Xiao

et al. 2009), whereas the negative dataset is often generated artificially and rarely based on negative miRNA–mRNA target results. The positive dataset can introduce bias caused by cell type, condition, and laboratory approaches (Schafer and Ciaudo 2020). The negative dataset can be too similar to the positive dataset or too artificial compared to real miRNA–mRNA non-targets. Both situations will lead to a ML approach not being able to efficiently discriminate between target and non-target (Riolo et al. 2020). As important as the creation of the positive and negative dataset are the balancing of both sets to prevent over or under-fitted models (Parveen et al. 2019). In summary, the ML-based miRNA target prediction methods use in principle the same miRNA–mRNA interaction features but differ mainly in the chosen ML architecture and the training dataset used for training and validation. For example, miTarget (SVM), miTAR (RF), and mirMark (NN) are all based on seed matching and different free energy features but differ in the data source and sample size. For example, miTarget is trained on 398 positive meaningful miRNA–mRNA target sites from the literature, mirMark is trained on the full miRecords database (Xiao et al. 2009) and miTAR used ~33,000 positive targets extracted from the DeepMirTar and miRAW training datasets.

4.3 Best Practices for miRNA Target Prediction

Due to the diverse types of miRNA target binding sites based on position and localization of Watson–Crick pairings and mismatches it is important to use several features and tools for the prediction (Bartel 2009). The interaction complexity allow only a limited generalizability of rules for the miRNA–mRNA target prediction (Witkos et al. 2011). Based on the heterogeneity of target sites, the usage of a single miRNA–mRNA target feature or model is not sufficient and requires specifically developed best practices and workflows.

4.3.1 Workflow to Detect miRNA–mRNA Target Sites

Target prediction programs are in principle binary classifiers and the two main statistical parameters to proof the tools performance are based on specificity and sensitivity. A comparative study of target prediction algorithms (Sethupathy et al. 2006) checked for these parameters using a set of experimentally validated mammalian targets from a database, which did not include all possible types of miRNA–mRNA target sites. Tools focusing on conserved seed matching, e.g., TargetScan (Lewis et al. 2003) reached a specificity of around 50% and sensitivity of 6–12% for specific miRNAs. Combination of several tools in so called metatools (Box 4.3) can increase the sensitivity at the cost of specificity creating a consensus out of tools considering different features like length of the 3′ UTR or synergistic effects from multiple target sites of the same or different miRNAs (Gaidatzis et al. 2007).

Box 4.3 Metatools of miRNA Target Prediction

Combining results from various tools can lead to a decrease in the prediction performance meaning the intersection of results from two or more tools can improve specificity at the cost of decreasing sensitivity, whereas the union of two or more tools increases the number of true targets and decreases the specificity (Witkos et al. 2011). However, in many studies, ensemble methods like the intersection of multiple tools are used to avoid false-positive predictions regardless of the loss in sensitivity (Wang et al. 2017). In such approaches, a ranking system for miRNA–mRNA is adapted based on the strength of the correlation coefficients (Le et al. 2015). For such a rank, the Borda count election can be used by selecting candidates in a democratic election with the best average rank (Marbach et al. 2012). Ensemble methods help with obtaining more comprehensive results, at which a combination of methods taking different approaches may result in a better ensemble method than combining methods in the same category. Nevertheless, simple meta-strategies like the intersection to directly integrate the outputs of individual predictors may not improve the prediction performance significantly (Zhao and Xue 2017). Therefore, more sophisticated meta-strategies can balance out sensitivity and specificity of different tools by integrating various data analysis techniques. Additionally, it can be expected that additional techniques may be used to further reduce noise, and consequently improve prediction accuracy but so far no consensus regarding the gold standard for miRNA target prediction exists (Oliveira et al. 2017).

In general, target prediction programs give scores and percentages to positions that only assess the possibility of interaction. By this, the basic concept of miRNA–mRNA target prediction algorithms predetermines the outcome advantages and weak points. For instance, the wobble pairing within the seed region of miRanda (Enright et al. 2003) adds 3′ compensatory sites by a simultaneous lowering of the precision. Another example would be DIANA-microT (Paraskevopoulou et al. 2013) analyzing target sites independently leading to a bias against miRNAs with multiple target sites. To overcome such shortcomings best practices for miRNA–mRNA target site prediction combining several tools are necessary and dependent on the wished outcome like performing comprehensive analysis to discover all true interactions (basic research) or finding the strongest interactions that could be employed in gene therapies (clinical-oriented research) (Witkos et al. 2011).

A possible best practices workflow (Fig. 4.4) would be to use in a first step one program focusing on site conservation like seed matching (e.g., MiRTif, miTAR, or TargetScan) because they are characterized by high precision and sensitivity. The next optional step is to add targets indicated by programs exploiting other parameters for final scoring (e.g., Cupid, TarPmiR, or DeepMirTar). After selection of predicted miRNA–mRNA target sites tools like ComiR, TargetMiner, TargetExpress, or Oasis 2.0 are using the expression profile of miRNAs and targets to detect overlaps or

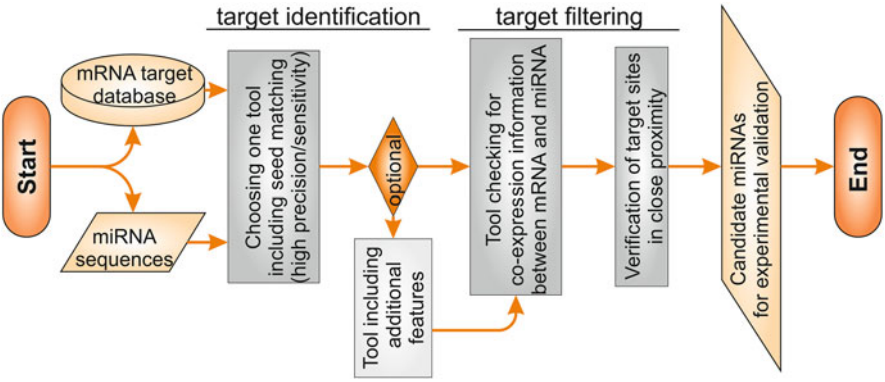


Fig. 4.4 Flowchart for best practices in miRNA target prediction. The flowchart shows a best course of practice to predict miRNA–mRNA targets. Input, Output (parallelogram), and database are orange, essential processing steps are gray and optional processing steps are light gray

inverse correlations. Finally, the last step to detect putative target sites in the immediate vicinity (MirTarget, miSTAR, or deepTarget) would be beneficial as it may play an important role in the mechanism of miRNA-mediated gene regulation. In genome-wide analyses (Grimson et al. 2007) as well as experimental data (Saetrom et al. 2007), it could be demonstrated that target sites located close to each other often act synergistically. After determining candidate miRNAs and their target mRNA interactions, synergy has to be validated using wet lab experiments to obtain full legitimacy (Thomson et al. 2011). Experimentally validated miRNA–mRNA interactions can be found in various databases like TarBase (Papadopoulos et al. 2009), MiRecords (Xiao et al. 2009), Ago (Shahi et al. 2006), or miRNAMAP (Hsu et al. 2008).

4.3.2 Validation of miRNA Target Prediction

Several methods for experimental verification of predicted miRNA–mRNA interactions are currently being used, i.e., reporter assays, transcriptomics, and proteomics analyses (Witkos et al. 2011). Depending on the information provided, experimental approaches can be classified into direct and indirect methods. Direct methods like reporter assays investigate the existence of an interaction between miRNA and its target, by studying directly the miRNA–mRNA pair or introducing a specific target site bound by miRNAs, known as MREs, into a reporter gene that measures the potential miRNA-induced changes at protein levels. Transcriptomics and proteomics analyses are categorized as indirect approaches because in both cases effects derived from an altered miRNA expression on mRNA or protein expression are observed. However, both classes have their disadvantages. Direct methods are relying on predictions of MREs and indirect methods can be affected by "knock-on" effects

altering indirectly the expression of multiple genes (Alexiou et al. 2009). An easy measurement approach under the gene reporter assays to detect miRNA–mRNA interactions are luciferase reporter assays (Miranda et al. 2006). This method is based on cloning the 3′ UTRs of genes (containing miRNA binding sites) into expression vectors bearing a reporter gene and using as negative control mutated target sites. Reporter assays serve as an efficient strategy for the verification of individual miRNA–mRNA interactions. In contrast, transcriptomics and proteomics approaches enable a genome-wide analysis of putative miRNA/target interactions. Microarray experiments compare cell transcriptomes after miRNA overexpression or inhibition with reference to the transcriptome of untreated cells which are highly dependent on the cell physiology. For proteome analysis, stable isotope labeling with amino acids in cell culture (SILAC) is used followed by a quantitative-mass spectrometry (Vinther et al. 2006). Besides these two main high-throughput methods approaches like AGO immunoprecipitation followed by NGS are used to identify miRNA–mRNA target prediction at a genome-wide level (Chi et al. 2009).

As a general approach, four criteria should be fulfilled to validate a miRNA–mRNA target pair in the biological model of interest (Thomson et al. 2011): (1) Co-expression of miRNA and predicted target mRNA must be demonstrated; (2) A direct interaction between the miRNA of interest and a specific region within the target mRNA must be proved; (3) Gain and loss of function experiments must be performed to demonstrate the regulatory mechanism; (4) The predicted changes in protein expression that are associated with modified biological functions should be demonstrated (Matkovich et al. 2011). First, the co-profiling of miRNAs and mRNAs allows insights into sharing the same transcriptional program or regulation by members of the same pathway. Microarray and RNA-Seq profiling on large scales as well as northern blots and RT-qPCR on a few genes can demonstrate co-expression. Second, it is essential to investigate the physical interaction between the miRNA and the MRE in the target mRNA. For such a direct interaction the reporter assays are still the gold standard procedure. The 3′ UTR with the MRE sequence is cloned downstream of the reporter gene and subcloned under the control of a ubiquitous promoter. Third, protein abundance analysis should be performed to investigate the change at the protein level of the target mRNA bound by the miRNA. Therefore, conventional methods like western blotting, ELISA or immune-cytochemistry experiments as well as global analysis like SILAC with additional mass spectrometry can be used (Yang et al. 2010). Fourth, it is essential to show that the miRNA–mRNA interaction has a biological function in the cell via in vitro or in vivo assays (Orom et al. 2008). By such assays signaling pathways, proliferation, differentiation, or migration behavior can be analyzed in cellular models.

As the experimental validation of miRNA targets is difficult, miRNA target prediction should be a powerful tool to identify potential miRNA targets. Computational approaches should provide invaluable tools for identification of miRNA targets in various biological networks. To accurately answer such questions more complex computational approaches are needed modeling the interplay between miRNAs (Huang et al. 2010). The present scarcity of experimentally validated exact miRNA hybridization sites is one big obstacle to the development of better

prediction methods (Riolo et al. 2020). High-throughput methods can on the one hand connect computational prediction and experimental validation but on the other hand generate a large amount of data that is not easy to interpret. Further, the validation of miRNA targets via such indirect methods can lead to "knock-on" effects by overexpression studies or are limited by the specificity of the inhibitor (Thomson et al. 2011).

4.4 miRNA-Based Therapeutics for Cancer

Worldwide, an estimated 19.3 million new cancer cases and almost 10.0 million cancer deaths occurred in 2020 (Sung et al. 2021). According to estimates from the World Health Organization (WHO) in 2019, cancer is the first or second leading cause of death before the age of 70 years in 112 of 183 countries and ranks third or fourth in a further 23 countries. Cancer's rising prominence as a leading cause of death partly reflects marked declines in mortality rates of stroke and coronary heart disease, relative to cancer, in many countries. At present, three main therapeutic strategies for cancer management are available (chemotherapy, surgery, and radiotherapy). Nowadays, direct and indirect miRNA-based strategies for cancer treatment have been developed (Box 4.4) with the potential to increase survival rate and reduce mortality rate (Gambari et al. 2016). Therapies are either based on reduction/inhibition of miRNAs or miRNAs are replaced/restored (Kong et al. 2012).

Box 4.4 Strategies for miRNA-Based Therapeutics

Blocking oncogenic miRNAs using antisense oligonucleotides—Antisense oligonucleotides are used as competitive inhibitors of miRNAs by annealing to the mature miRNA guide strand and inducing degradation or stoichiometric duplex formation (Hutvagner et al. 2004).

Locked nucleic acid (LNA) antimiR constructs—Nucleic acid analogues like LNA nucleosides contain a "locked" ribose ring by a methylene bridge connecting the 2′-O atom and the 4′-C atom. By "locking" the molecule with the methylene bridge, LNA oligonucleotides increase mismatches discrimination, aqueous solubility, and hybridization affinity (Vester and Wengel 2004).

miRNAs sponges—miRNA sponges are decoy transcripts containing multiple, tandem binding sites to a miRNA of interest. Beside antisense oligonucleotides against miRNAs (Ebert et al. 2007).

miR-Mask—miR-Mask stands for miRNA-Masking antisense oligonucleotides technology and is another decoy-based mechanism. miR-Masks consist of single-stranded 2′-*O*-methyl-modified antisense oligonucleotides that are fully complementary to predicted miRNA binding sites in the 3′ UTR of the target mRNA (Choi et al. 2007).

(continued)

Box 4.4 (continued)

Small-molecule inhibitors—Small molecule drugs are used to modulate the expression of miRNAs by targeting signaling pathways. This will influence the activation of transcription factors that regulate miRNA encoding genes. The aim is to modulate the miRNA maturation and degradation process machinery (Gumireddy et al. 2008).

Restoring tumor suppressor miRNA expression—Compensation for the loss or downregulation of a tumor suppressor miRNA can be reached by the introduction of synthetic oligonucleotides like miRNA mimics (Garzon et al. 2009).

Reprogramming cancer cells—In contrast, this strategy is based on reprogramming a miRNA network in cancer. The reprogramming could be achieved by chemotherapeutic drugs or the modulation of several antisense oligonucleotides or miRNA mimics (Garzon et al. 2007).

Many promising examples potentially lead to the development of miRNA-based therapeutic protocols against cancer, in which therapies should be designed to target multiple miRNAs in respect to the high complexity of the networks for cancer (Gambari et al. 2016). The problem of in vivo miRNA therapeutics is the delivery to specific tissues and of sufficient amount of cellular uptake due to biological instability of unmodified oligonucleotides, size, and negative charge (Aagaard and Rossi 2007). In addition, the potential off-target effects of miRNA therapeutics may cause toxic phenotypes (Dias and Stein 2002). But besides the importance of miRNAs as a therapeutic possibility for cancer miRNAs have also a very important implication in the diagnosis and prognosis of cancer types. One relatively recent example is the specific pattern of circulating cell-free miRNAs in serum (Fayyad-Kazan et al. 2013). Such cancer-specific miRNAs present in extracellular body fluids allow the crosstalk between cancer and normal cells and has the advantage for diagnosis and prognosis to be a noninvasive liquid biopsy (Koberle et al. 2013).

References

Aagaard L, Rossi JJ (2007) RNAi therapeutics: principles, prospects and challenges. Adv Drug Deliv Rev 59(2–3):75–86. https://doi.org/10.1016/j.addr.2007.03.005

Adams L (2017) Pri-miRNA processing: structure is key. Nat Rev Genet 18(3):145–145. https://doi.org/10.1038/nrg.2017.6

Ahmadi H, Ahmadi A, Azimzadeh-Jamalkandi S, Shoorehdeli MA, Salehzadeh-Yazdi A, Bidkhori G, Masoudi-Nejad A (2013) HomoTarget: a new algorithm for prediction of microRNA targets in Homo sapiens. Genomics 101(2):94–100. https://doi.org/10.1016/j.ygeno.2012.11.005

Alexiou P, Maragkakis M, Papadopoulos GL, Reczko M, Hatzigeorgiou AG (2009) Lost in translation: an assessment and perspective for computational microRNA target identification. Bioinformatics 25(23):3049–3055. https://doi.org/10.1093/bioinformatics/btp565

Bandyopadhyay S, Mitra R (2009) TargetMiner: microRNA target prediction with systematic identification of tissue-specific negative examples. Bioinformatics 25(20):2625–2631. https://doi.org/10.1093/bioinformatics/btp503

Bandyopadhyay S, Ghosh D, Mitra R, Zhao ZM (2015) MBSTAR: multiple instance learning for predicting specific functional binding sites in microRNA targets. Sci Rep-UK 5. https://doi.org/10.1038/srep08004

Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136(2):215–233. https://doi.org/10.1016/j.cell.2009.01.002

Benway CJ, Iacomini J (2018) Defining a microRNA-mRNA interaction map for calcineurin inhibitor induced nephrotoxicity. Am J Transplant 18(4):796–809. https://doi.org/10.1111/ajt.14503

Betel D, Koppal A, Agius P, Sander C, Leslie C (2010) Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. Genome Biol 11(8):R90. https://doi.org/10.1186/gb-2010-11-8-r90

Capece V, Vizcaino JCG, Vidal R, Rahman RU, Centeno TP, Shomroni O, Suberviola I, Fischer A, Bonn S (2015) Oasis: online analysis of small RNA deep sequencing data. Bioinformatics 31(13):2205–2207. https://doi.org/10.1093/bioinformatics/btv113

Chakraborty C, Sharma AR, Sharma G, Doss CGP, Lee SS (2017) Therapeutic miRNA and siRNA: moving from bench to clinic as next generation medicine. Mol Ther Nucleic Acids 8:132–143. https://doi.org/10.1016/j.omtn.2017.06.005

Chakraborty C, Sharma AR, Sharma G, Sarkar BK, Lee SS (2018) The novel strategies for next-generation cancer treatment: miRNA combined with chemotherapeutic agents for the treatment of cancer. Oncotarget 9(11):10164–10174. https://doi.org/10.18632/oncotarget.24309

Chandra V, Girijadevi R, Nair AS, Pillai SS, Pillai RM (2010) MTar: a computational microRNA target prediction architecture for human transcriptome. Bmc Bioinformatics 11. https://doi.org/10.1186/1471-2105-11-S1-S2

Chang DT, Wang CC, Chen JW (2008) Using a kernel density estimation based classifier to predict species-specific microRNA precursors. BMC Bioinformatics 9(Suppl 12):S2. https://doi.org/10.1186/1471-2105-9-S12-S2

Cheng S, Guo MZ, Wang CY, Liu XY, Liu Y, Wu XJ (2016) MiRTDL: a deep learning approach for miRNA target prediction. IEEE Acm T Comput Bi 13(6):1161–1169. https://doi.org/10.1109/Tcbb.2015.2510002

Chi SW, Zang JB, Mele A, Darnell RB (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature 460(7254):479–486. https://doi.org/10.1038/nature08170

Chipman LB, Pasquinelli AE (2019) miRNA targeting: growing beyond the seed. Trends Genet 35(3):215–222. https://doi.org/10.1016/j.tig.2018.12.005

Chiu HS, Llobet-Navas D, Yang X, Chung WJ, Ambesi-Impiombato A, Iyer A, Kim HR, Seviour EG, Luo Z, Sehgal V, Moss T, Lu Y, Ram P, Silva J, Mills GB, Califano A, Sumazin P (2015) Cupid: simultaneous reconstruction of microRNA-target and ceRNA networks. Genome Res 25(2):257–267. https://doi.org/10.1101/gr.178194.114

Choi WY, Giraldez AJ, Schier AF (2007) Target protectors reveal dampening and balancing of nodal agonist and antagonist by miR-430. Science 318(5848):271–274. https://doi.org/10.1126/science.1147535

Chu YW, Chang KP, Chen CW, Liang YT, Soh ZT, Hsieh LC (2020) miRgo: integrating various off-the-shelf tools for identification of microRNA-target interactions by heterogeneous features and a novel evaluation indicator. Sci Rep 10(1):1466. https://doi.org/10.1038/s41598-020-58336-5

Dias N, Stein CA (2002) Antisense oligonucleotides: basic concepts and mechanisms. Mol Cancer Ther 1(5):347–355

Ding J, Li XM, Hu HY (2016) TarPmiR: a new approach for microRNA target site prediction. Bioinformatics 32(18):2768–2775. https://doi.org/10.1093/bioinformatics/btw318

Ebert MS, Neilson JR, Sharp PA (2007) MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. Nat Methods 4(9):721–726. https://doi.org/10.1038/Nmeth1079

Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS (2003) MicroRNA targets in drosophila. Genome Biol 5(1):R1. https://doi.org/10.1186/gb-2003-5-1-r1

Fayyad-Kazan H, Bitar N, Najar M, Lewalle P, Fayyad-Kazan M, Badran R, Hamade E, Daher A, Hussein N, ElDirani R, Berri F, Vanhamme L, Burny A, Martiat P, Rouas R, Badran B (2013) Circulating miR-150 and miR-342 in plasma are novel potential biomarkers for acute myeloid leukemia. J Transl Med 11. https://doi.org/10.1186/1479-5876-11-31

Fernandez S, Risolino M, Mandia N, Talotta F, Soini Y, Incoronato M, Condorelli G, Banfi S, Verde P (2015) miR-340 inhibits tumor cell proliferation and induces apoptosis by targeting multiple negative regulators of p27 in non-small cell lung cancer. Oncogene 34(25):3240–3250. https://doi.org/10.1038/onc.2014.267

Gaidatzis D, van Nimwegen E, Hausser J, Zavolan M (2007) Inference of miRNA targets using evolutionary conservation and pathway analysis. Bmc Bioinformatics 8. https://doi.org/10.1186/1471-2105-8-69

Gambari R, Brognara E, Spandidos DA, Fabbri E (2016) Targeting oncomiRNAs and mimicking tumor suppressor miRNAs: Nuew trends in the development of miRNA therapeutic strategies in oncology (review). Int J Oncol 49(1):5–32. https://doi.org/10.3892/ijo.2016.3503

Garcia DM, Baek D, Shin C, Bell GW, Grimson A, Bartel DP (2011) Weak seed-pairing stability and high target-site abundance decrease the proficiency of lsy-6 and other microRNAs. Nat Struct Mol Biol 18(10):1139–1146. https://doi.org/10.1038/nsmb.2115

Garzon R, Pichiorri F, Palumbo T, Visentini M, Aqeilan R, Cimmino A, Wang H, Sun H, Volinia S, Alder H, Calin GA, Liu CG, Andreeff M, Croce CM (2007) MicroRNA gene expression during retinoic acid-induced differentiation of human acute promyelocytic leukemia. Oncogene 26(28):4148–4157. https://doi.org/10.1038/sj.onc.1210186

Garzon R, Heaphy CE, Havelange V, Fabbri M, Volinia S, Tsao T, Zanesi N, Kornblau SM, Marcucci G, Calin GA, Andreeff M, Croce CM (2009) MicroRNA 29b functions in acute myeloid leukemia. Blood 114(26):5331–5341. https://doi.org/10.1182/blood-2009-03-211938

Garzon R, Marcucci G, Croce CM (2010) Targeting microRNAs in cancer: rationale, strategies and challenges. Nat Rev Drug Discov 9(10):775–789. https://doi.org/10.1038/nrd3179

Ghosal S, Saha S, Das S, Sen R, Goswami S, Jana SS, Chakrabarti J (2016) miRepress: modelling gene expression regulation by microRNA with non-conventional binding sites. Sci Rep-UK 6. https://doi.org/10.1038/srep22334

Ghoshal A, Shankar R, Bagchi S, Grama A, Chaterji S (2015) MicroRNA target prediction using thermodynamic and sequence curves. BMC Genomics 16:999. https://doi.org/10.1186/s12864-015-1933-2

Ghoshal A, Shankar R, Bagchi S, Grama A, Chaterji S (2016) MicroRNA target prediction using thermodynamic and sequence curves (vol 16, 999, 2015). BMC Genomics 17. https://doi.org/10.1186/s12864-016-2367-1

Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR (2003) Rfam: an RNA family database. Nucleic Acids Res 31(1):439–441. https://doi.org/10.1093/nar/gkg006

Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP (2007) MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell 27(1):91–105. https://doi.org/10.1016/j.molcel.2007.06.017

Gruber AR, Lorenz R, Bernhart SH, Neubock R, Hofacker IL (2008) The Vienna RNA websuite. Nucleic Acids Res 36(Web Server issue):W70–W74. https://doi.org/10.1093/nar/gkn188

Gu T, Zhao X, Barbazuk WB, Lee JH (2021) miTAR: a hybrid deep learning-based approach for predicting miRNA targets. BMC Bioinformatics 22(1):96. https://doi.org/10.1186/s12859-021-04026-6

Gumireddy K, Young DD, Xiong X, Hogenesch JB, Huang QH, Deiters A (2008) Small-molecule inhibitors of microRNA miR-21 function. Angew Chem Int Edit 47(39):7482–7484. https://doi.org/10.1002/anie.200801555

Hsu SD, Chu CH, Tsou AP, Chen SJ, Chen HC, Hsu PW, Wong YH, Chen YH, Chen GH, Huang HD (2008) miRNAMap 2.0: genomic maps of microRNAs in metazoan genomes. Nucleic Acids Res 36(Database issue):D165–D169. https://doi.org/10.1093/nar/gkm1012

Hsu SD, Lin FM, Wu WY, Liang C, Huang WC, Chan WL, Tsai WT, Chen GZ, Lee CJ, Chiu CM, Chien CH, Wu MC, Huang CY, Tsou AP, Huang HD (2011) miRTarBase: a database curates experimentally validated microRNA-target interactions. Nucleic Acids Res 39(Database issue): D163–D169. https://doi.org/10.1093/nar/gkq1107

Huang JC, Babak T, Corson TW, Chua G, Khan S, Gallie BL, Hughes TR, Blencowe BJ, Frey BJ, Morris QD (2007) Using expression profiling data to identify human microRNA targets. Nat Methods 4(12):1045–1049. https://doi.org/10.1038/nmeth1130

Huang Y, Zou Q, Song HT, Song F, Wang LG, Zhang GZ, Shen XJ (2010) A study of miRNAs targets prediction and experimental validation. Protein Cell 1(11):979–986. https://doi.org/10.1007/s13238-010-0129-4

Huttenhofer A, Brosius J, Bachellerie JP (2002) RNomics: identification and function of small, non-messenger RNAs. Curr Opin Chem Biol 6(6):835–843. https://doi.org/10.1016/s1367-5931(02)00397-6

Hutvagner G, Simard MJ, Mello CC, Zamore PD (2004) Sequence-specific inhibition of small RNA function. PLoS Biol 2(4):E98. https://doi.org/10.1371/journal.pbio.0020098

Jiang P, Wu H, Wang W, Ma W, Sun X, Lu Z (2007) MiPred: classification of real and pseudo microRNA precursors using random forest prediction model with combined features. Nucleic Acids Res 35(Web Server issue):W339–W344. https://doi.org/10.1093/nar/gkm368

Kadri S, Hinman V, Benos PV (2009) HHMMiR: efficient de novo prediction of microRNAs using hierarchical hidden Markov models. BMC Bioinformatics 10(Suppl 1):S35. https://doi.org/10.1186/1471-2105-10-S1-S35

Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I, Papadimitriou D, Kavakiotis I, Maniou S, Skoufos G, Vergoulis T, Dalamagas T, Hatzigeorgiou AG (2018) DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. Nucleic Acids Res 46(D1):D239–D245. https://doi.org/10.1093/nar/gkx1141

Kern F, Backes C, Hirsch P, Fehlmann T, Hart M, Meese E, Keller A (2020) What's the target: understanding two decades of in silico microRNA-target prediction. Brief Bioinform 21(6): 1999–2010. https://doi.org/10.1093/bib/bbz111

Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E (2007) The role of site accessibility in microRNA target recognition. Nat Genet 39(10):1278–1284. https://doi.org/10.1038/ng2135

Kim SK, Nam JW, Rhee JK, Lee WJ, Zhang BT (2006) miTarget: microRNA target gene prediction using a support vector machine. Bmc Bioinformatics 7. https://doi.org/10.1186/1471-2105-7-411

Kishore S, Jaskiewicz L, Burger L, Hausser J, Khorshid M, Zavolan M (2011) A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. Nat Methods 8(7):559–U561. https://doi.org/10.1038/Nmeth.1608

Koberle V, Kronenberger B, Pleli T, Trojan J, Imelmann E, Peveling-Oberhag J, Welker MW, Elhendawy M, Zeuzem S, Piiper A, Waidmann O (2013) Serum microRNA-1 and microRNA-122 are prognostic markers in patients with hepatocellular carcinoma. Eur J Cancer 49(16): 3442–3449. https://doi.org/10.1016/j.ejca.2013.06.002

Kong YW, Ferland-McCollough D, Jackson TJ, Bushell M (2012) microRNAs in cancer management. Lancet Oncol 13(6):e249–e258. https://doi.org/10.1016/S1470-2045(12)70073-6

Kozomara A, Birgaoanu M, Griffiths-Jones S (2019) miRBase: from microRNA sequences to function. Nucleic Acids Res 47(D1):D155–D162. https://doi.org/10.1093/nar/gky1141

Le TD, Zhang J, Liu L, Li J (2015) Ensemble methods for MiRNA target prediction from expression data. PLoS One 10(6):e0131627. https://doi.org/10.1371/journal.pone.0131627

Leclercq M, Diallo AB, Blanchette M (2017) Prediction of human miRNA target genes using computationally reconstructed ancestral mammalian sequences. Nucleic Acids Res 45(2): 556–566. https://doi.org/10.1093/nar/gkw1085

Lee B, Baek J, Park S, Yoon S (2016) deepTarget: end-to-end learning framework for microRNA target prediction using deep recurrent neural networks. In: Proceedings of the 7th ACM international conference on bioinformatics, computational biology, and health informatics, pp 434–442. https://doi.org/10.1145/2975167.2975212

Lekprasert P, Mayhew M, Ohler U (2011) Assessing the utility of thermodynamic features for microRNA target prediction under relaxed seed and no conservation requirements. PLoS One 6(6). https://doi.org/10.1371/journal.pone.0020622

Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of mammalian microRNA targets. Cell 115(7):787–798. https://doi.org/10.1016/s0092-8674(03)01018-3

Liu W, Wang X (2019) Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. Genome Biol 20(1):18. https://doi.org/10.1186/s13059-019-1629-z

Liu H, Yue D, Chen Y, Gao SJ, Huang Y (2010a) Improving performance of mammalian microRNA target prediction. BMC Bioinformatics 11:476. https://doi.org/10.1186/1471-2105-11-476

Liu H, Yue D, Zhang L, Chen Y, Gao SJ, Huang Y (2010b) A Bayesian approach for identifying miRNA targets by combining sequence prediction and gene expression profiling. BMC Genomics 11(Suppl 3):S12. https://doi.org/10.1186/1471-2164-11-S3-S12

Liu H, Zhou SGN, Guan JH (2013) Identifying mammalian MicroRNA targets based on supervised distance metric learning. IEEE J Biomed Health 17(2):427–435. https://doi.org/10.1109/Titb.2012.2229286

Liu Y, Luo JW, Ding PJ (2019) Inferring MicroRNA targets based on restricted Boltzmann machines. IEEE J Biomed Health 23(1):427–436. https://doi.org/10.1109/Jbhi.2018.2814609

Lu Y, Leslie CS (2016) Learning to predict miRNA-mRNA interactions from AGO CLIP sequencing and CLASH data. PLoS Comput Biol 12(7):e1005026. https://doi.org/10.1371/journal.pcbi.1005026

Lukasik A, Wojcikowski M, Zielenkiewicz P (2016) Tools4miRs - one place to gather all the tools for miRNA analysis. Bioinformatics 32(17):2722–2724. https://doi.org/10.1093/bioinformatics/btw189

Lytle JR, Yario TA, Steitz JA (2007) Target mRNAs are repressed as efficiently by microRNA-binding sites in the 5′ UTR as in the 3′ UTR. Proc Natl Acad Sci U S A 104(23):9667–9672. https://doi.org/10.1073/pnas.0703820104

Marbach D, Costello JC, Kuffner R, Vega NM, Prill RJ, Camacho DM, Allison KR, Consortium D, Kellis M, Collins JJ, Stolovitzky G (2012) Wisdom of crowds for robust gene network inference. Nat Methods 9(8):796–804. https://doi.org/10.1038/nmeth.2016

Marin RM, Vanicek J (2011) Efficient use of accessibility in microRNA target prediction. Nucleic Acids Res 39(1):19–29. https://doi.org/10.1093/nar/gkq768

Matkovich SJ, Van Booven DJ, Eschenbacher WH, Dorn GW 2nd (2011) RISC RNA sequencing for context-specific identification of in vivo microRNA targets. Circ Res 108(1):18–26. https://doi.org/10.1161/CIRCRESAHA.110.233528

Menor M, Ching T, Zhu X, Garmire D, Garmire LX (2014) mirMark: a site-level and UTR-level classifier for miRNA target prediction. Genome Biol 15(10). https://doi.org/10.1186/s13059-014-0500-5

Min H, Yoon S (2010) Got target? Computational methods for microRNA target prediction and their extension. Exp Mol Med 42(4):233–244. https://doi.org/10.3858/emm.2010.42.4.032

Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, Lim B, Rigoutsos I (2006) A pattern-based method for the identification of microRNA binding sites and their corresponding heteroduplexes. Cell 126(6):1203–1217. https://doi.org/10.1016/j.cell.2006.07.031

Misiewicz-Krzeminska I, Krzeminski P, Corchete LA, Quwaider D, Rojas EA, Herrero AB, Gutierrez NC (2019) Factors regulating microRNA expression and function in multiple myeloma. Noncoding RNA 5(1):doi:10.3390/ncrna5010009

Mitra R, Bandyopadhyay S (2011) MultiMiTar: a novel multi objective optimization based miRNA-target prediction method. PLoS One 6(9):e24583. https://doi.org/10.1371/journal.pone.0024583

Muniategui A, Nogales-Cadenas R, Vazquez M, Aranguren XL, Agirre X, Luttun A, Prosper F, Pascual-Montano A, Rubio A (2012) Quantification of miRNA-mRNA interactions. PLoS One 7(2):10.1371/journal.pone.0030766

Nam JW, Kim J, Kim SK, Zhang BT (2006) ProMiR II: a web server for the probabilistic prediction of clustered, nonclustered, conserved and nonconserved microRNAs. Nucleic Acids Res 34 (Web Server issue):W455–W458. https://doi.org/10.1093/nar/gkl321

Oliveira AC, Bovolenta LA, Nachtigall PG, Herkenhoff ME, Lemke N, Pinhal D (2017) Combining results from distinct MicroRNA target prediction tools enhances the performance of analyses. Front Genet 8. https://doi.org/10.3389/fgene.2017.00059

Orom UA, Nielsen FC, Lund AH (2008) MicroRNA-10a binds the 5′ UTR of ribosomal protein mRNAs and enhances their translation. Mol Cell 30(4):460–471. https://doi.org/10.1016/j.molcel.2008.05.001

Ovando-Vazquez C, Lepe-Soltero D, Abreu-Goodger C (2016) Improving microRNA target prediction with gene expression profiles. BMC Genomics 17:364. https://doi.org/10.1186/s12864-016-2695-1

Papadopoulos GL, Reczko M, Simossis VA, Sethupathy P, Hatzigeorgiou AG (2009) The database of experimentally supported targets: a functional update of TarBase. Nucleic Acids Res 37 (Database issue):D155–D158. https://doi.org/10.1093/nar/gkn809

Paraskevopoulou MD, Georgakilas G, Kostoulas N, Vlachos IS, Vergoulis T, Reczko M, Filippidis C, Dalamagas T, Hatzigeorgiou AG (2013) DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. Nucleic Acids Res 41(W1):W169–W173. https://doi.org/10.1093/nar/gkt393

Paraskevopoulou MD, Karagkouni D, Vlachos IS, Tastsoglou S, Hatzigeorgiou AG (2018) microCLIP super learning framework uncovers functional transcriptome-wide miRNA interactions. Nat Commun 9. https://doi.org/10.1038/s41467-018-06046-y

Parveen A, Mustafa SH, Yadav P, Kumar A (2019) Applications of machine learning in miRNA discovery and target prediction. Curr Genomics 20(8):537–544. https://doi.org/10.2174/1389202921666200106111813

Pelaez N, Carthew RW (2012) Biological robustness and the role of microRNAs: a network perspective. Curr Top Dev Biol 99:237–255. https://doi.org/10.1016/B978-0-12-387038-4.00009-4

Peterson SM, Thompson JA, Ufkin ML, Sathyanarayana P, Liaw L, Congdon CB (2014) Common features of microRNA target prediction tools. Front Genet 5:23. https://doi.org/10.3389/fgene.2014.00023

Pla A, Zhong XF, Rayner S (2018) miRAW: a deep learning-based approach to predict microRNA targets by analyzing whole microRNA transcripts. PLoS Comput Biol 14(7):10.1371/journal.pcbi.1006185

Rennie W, Liu CC, Carmack CS, Wolenc A, Kanoria S, Lu J, Long D, Ding Y (2014) STarMir: a web server for prediction of microRNA binding sites. Nucleic Acids Res 42(W1):W114–W118. https://doi.org/10.1093/nar/gku376

Reyes-Herrera PH, Ficarra E, Acquaviva A, Macii E (2011) miREE: miRNA recognition elements ensemble. BMC Bioinformatics 12:454. https://doi.org/10.1186/1471-2105-12-454

Riolo G, Cantara S, Marzocchi C, Ricci C (2020) miRNA targets: from prediction tools to experimental validation. Methods Protoc 4(1):10.3390/mps4010001

Rojo Arias JE, Busskamp V (2019) Challenges in microRNAs' targetome prediction and validation. Neural Regen Res 14(10):1672–1677. https://doi.org/10.4103/1673-5374.257514

Rumelhart DE, Hinton GE, Williams RJ (1986) Learning representations by back-propagating errors. Nature 323(6088):533–536. https://doi.org/10.1038/323533a0

Saetrom P, Heale BSE, Snove O, Aagaard L, Alluin J, Rossi JJ (2007) Distance constraints between microRNA target sites dictate efficacy and cooperativity. Nucleic Acids Res 35(7):2333–2342. https://doi.org/10.1093/nar/gkm133

Schafer M, Ciaudo C (2020) Prediction of the miRNA interactome – established methods and upcoming perspectives. Comput Struct Biotechnol J 18:548–557. https://doi.org/10.1016/j.csbj.2020.02.019

Sebastian B, Aggrey SE (2008) Specificity and sensitivity of PROMIR, ERPIN and MIR-ABELA in predicting pre-microRNAs in the chicken genome. In Silico Biol 8(5–6):377–381

Selbach M, Schwanhausser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N (2008) Widespread changes in protein synthesis induced by microRNAs. Nature 455(7209):58–63. https://doi.org/10.1038/nature07228

Sethupathy P, Megraw M, Hatzigeorgiou AG (2006) A guide through present computational approaches for the identification of mammalian microRNA targets. Nat Methods 3(11): 881–886. https://doi.org/10.1038/Nmeth954

Shahi P, Loukianiouk S, Bohne-Lang A, Kenzelmann M, Kuffer S, Maertens S, Eils R, Grone HJ, Gretz N, Brors B (2006) Argonaute--a database for gene regulation by mammalian microRNAs. Nucleic Acids Res 34(Database issue):D115–D118. https://doi.org/10.1093/nar/gkj093

Sidey-Gibbons JAM, Sidey-Gibbons CJ (2019) Machine learning in medicine: a practical introduction. BMC Med Res Methodol 19(1):64. https://doi.org/10.1186/s12874-019-0681-4

Simkin A, Geissler R, McIntyre ABR, Grimson A (2020) Evolutionary dynamics of microRNA target sites across vertebrate evolution. PLoS Genet 16(2):e1008285. https://doi.org/10.1371/journal.pgen.1008285

Siu MK, Tsai YC, Chang YS, Yin JJ, Suau F, Chen WY, Liu YN (2015) Transforming growth factor-beta promotes prostate bone metastasis through induction of microRNA-96 and activation of the mTOR pathway. Oncogene 34(36):4767–4776. https://doi.org/10.1038/onc.2014.414

Sturm M, Hackenberg M, Langenberger D, Frishman D (2010) TargetSpy: a supervised machine learning approach for microRNA target prediction. Bmc Bioinformatics 11. https://doi.org/10.1186/1471-2105-11-292

Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F (2021) Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. https://doi.org/10.3322/caac.21660

Svetnik V, Liaw A, Tong C, Culberson JC, Sheridan RP, Feuston BP (2003) Random forest: a classification and regression tool for compound classification and QSAR modeling. J Chem Inf Comput Sci 43(6):1947–1958. https://doi.org/10.1021/ci034160g

Thomson DW, Bracken CP, Goodall GJ (2011) Experimental strategies for microRNA target identification. Nucleic Acids Res 39(16):6845–6853. https://doi.org/10.1093/nar/gkr330

Van Peer G, De Paepe A, Stock M, Anckaert J, Volders PJ, Vandesompele J, De Baets B, Waegeman W (2017) miSTAR: miRNA target prediction through modeling quantitative and qualitative miRNA binding site information in a stacked model structure. Nucleic Acids Res 45(7):10.1093/nar/gkw1260

Vapnik VN (1999) An overview of statistical learning theory. IEEE T Neural Networ 10(5): 988–999. https://doi.org/10.1109/72.788640

Vester B, Wengel J (2004) LNA (locked nucleic acid): high-affinity targeting of complementary RNA and DNA. Biochemistry 43(42):13233–13241. https://doi.org/10.1021/bi0485732

Vinther J, Hedegaard MM, Gardner PP, Andersen JS, Arctander P (2006) Identification of miRNA targets with stable isotope labeling by amino acids in cell culture. Nucleic Acids Res 34(16): e107. https://doi.org/10.1093/nar/gkl590

Voinnet O (2009) Origin, biogenesis, and activity of plant microRNAs. Cell 136(4):669–687. https://doi.org/10.1016/j.cell.2009.01.046

Volinia S, Calin GA, Liu CG, Ambs S, Cimmino A, Petrocca F, Visone R, Iorio M, Roldo C, Ferracin M, Prueitt RL, Yanaihara N, Lanza G, Scarpa A, Vecchione A, Negrini M, Harris CC, Croce CM (2006) A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci U S A 103(7):2257–2261. https://doi.org/10.1073/pnas.0510565103

Wang XH, Cai P, Wang MH, Wang Z (2014) microRNA25 promotes osteosarcoma cell proliferation by targeting the cellcycle inhibitor p27. Mol Med Rep 10(2):855–859. https://doi.org/10.3892/mmr.2014.2260

Wang ZX, Xu WL, Liu Y (2015) Integrating full spectrum of sequence features into predicting functional microRNA-mRNA interactions. Bioinformatics 31(21):3529–3536. https://doi.org/10.1093/bioinformatics/btv392

Wang J, Liu H, Tian L, Wang F, Han L, Zhang W, Bai YA (2017) miR-15b inhibits the progression of glioblastoma cells through targeting insulin-like growth factor receptor 1. Horm Cancer 8(1): 49–57. https://doi.org/10.1007/s12672-016-0276-z

Weber JA, Baxter DH, Zhang S, Huang DY, Huang KH, Lee MJ, Galas DJ, Wang K (2010) The microRNA spectrum in 12 body fluids. Clin Chem 56(11):1733–1741. https://doi.org/10.1373/clinchem.2010.147405

Wen JY, Parker BJ, Jacobsen A, Krogh A (2011) MicroRNA transfection and AGO-bound CLIP-seq data sets reveal distinct determinants of miRNA action. RNA 17(5):820–834. https://doi.org/10.1261/rna.2387911

Wen M, Cong PS, Zhang ZM, Lu HM, Li TH (2018) DeepMirTar: a deep-learning approach for predicting human miRNA targets. Bioinformatics 34(22):3781–3787. https://doi.org/10.1093/bioinformatics/bty424

White NM, Fatoohi E, Metias M, Jung K, Stephan C, Yousef GM (2011) Metastamirs: a stepping stone towards improved cancer management. Nat Rev Clin Oncol 8(2):75–84. https://doi.org/10.1038/nrclinonc.2010.173

Witkos TM, Koscianska E, Krzyzosiak WJ (2011) Practical aspects of microRNA target prediction. Curr Mol Med 11(2):93–109. https://doi.org/10.2174/156652411794859250

Xia X, Li Y, Wang W, Tang F, Tan J, Sun L, Li Q, Sun L, Tang B, He S (2015) MicroRNA-1908 functions as a glioblastoma oncogene by suppressing PTEN tumor suppressor pathway. Mol Cancer 14:154. https://doi.org/10.1186/s12943-015-0423-0

Xiao F, Zuo Z, Cai G, Kang S, Gao X, Li T (2009) miRecords: an integrated resource for microRNA-target interactions. Nucleic Acids Res 37(Database issue):D105–D110. https://doi.org/10.1093/nar/gkn851

Xue C, Li F, He T, Liu GP, Li Y, Zhang X (2005) Classification of real and pseudo microRNA precursors using local structure-sequence features and support vector machine. BMC Bioinformatics 6:310. https://doi.org/10.1186/1471-2105-6-310

Yang Y, Wang YP, Li KB (2008) MiRTif: a support vector machine-based microRNA target interaction filter. BMC Bioinformatics 9(Suppl 12):S4. https://doi.org/10.1186/1471-2105-9-S12-S4

Yang Y, Chaerkady R, Kandasamy K, Huang TC, Selvan LDN, Dwivedi SB, Kent OA, Mendell JT, Pandey A (2010) Identifying targets of miR-143 using a SILAC-based proteomic approach. Mol BioSyst 6(10):1873–1882. https://doi.org/10.1039/c004401f

Yousef M, Nebozhyn M, Shatkay H, Kanterakis S, Showe LC, Showe MK (2006) Combining multi-species genomic data for microRNA identification using a Naive Bayes classifier. Bioinformatics 22(11):1325–1334. https://doi.org/10.1093/bioinformatics/btl094

Yousef M, Jung S, Kossenkov AV, Showe LC, Showe MK (2007) Naive Bayes for microRNA target predictions--machine learning for microRNA targets. Bioinformatics 23(22):2987–2992. https://doi.org/10.1093/bioinformatics/btm484

Yue D, Guo MZ, Chen YD, Huang YF (2012) A Bayesian decision fusion approach for microRNA target prediction. BMC Genomics 13. https://doi.org/10.1186/1471-2164-13-S8-S13

Zhao B, Xue B (2017) Improving prediction accuracy using decision-tree-based meta-strategy and multi-threshold sequential-voting exemplified by miRNA target prediction. Genomics 109(3–4):227–232. https://doi.org/10.1016/j.ygeno.2017.04.003

Zhao B, Xue B (2019) Significant improvement of miRNA target prediction accuracy in large datasets using meta-strategy based on comprehensive voting and artificial neural networks. BMC Genomics 20. https://doi.org/10.1186/s12864-019-5528-1

Zheng X, Fu X, Wang K, Wang M (2020) Deep neural networks for human microRNA precursor detection. BMC Bioinformatics 21(1):17. https://doi.org/10.1186/s12859-020-3339-7

Chapter 5
Turning Data to Knowledge: Online Tools, Databases, and Resources in microRNA Research

Ido Blass, Keren Zohar, and Michal Linial

Abstract MicroRNAs (miRNAs) provide a fundamental layer of regulation in cells. miRNAs act posttranscriptionally through complementary base-pairing with the 3′-UTR of a target mRNA, leading to mRNA degradation and translation arrest. The likelihood of forming a valid miRNA-target duplex within cells was computationally predicted and experimentally monitored. In human cells, the miRNA profiles determine their identity and physiology. Therefore, alterations in the composition of miRNAs signify many cancer types and chronic diseases. In this chapter, we introduce online functional tools and resources to facilitate miRNA research. We start by introducing currently available miRNA catalogs and miRNA-gateway portals for navigating among different miRNA-centric online resources. We then sketch several realistic challenges that may occur while investigating miRNA regulation in living cells. As a showcase, we demonstrate the utility of miRNAs and mRNAs expression databases that cover diverse human cells and tissues, including resources that report on genetic alterations affecting miRNA expression levels and alteration in binding capacity. Introducing tools linking miRNAs with transcription factor (TF) networks reveals miRNA regulation complexity within living cells. Finally, we concentrate on online resources that analyze miRNAs in human diseases and specifically in cancer. Altogether, we introduce contemporary, selected resources and online tools for studying miRNA regulation in cells and tissues and their utility in health and disease.

Keywords miRNA expression · Data retrieval · Data integration · Data mining · miRNA families · Cell states · ceRNA

I. Blass
The Rachel and Selim Benin School of Computer Science and Engineering, The Hebrew University of Jerusalem, Jerusalem, Israel

K. Zohar · M. Linial (✉)
Department of Biological Chemistry, Institute of Life Sciences, The Hebrew University of Jerusalem, Jerusalem, Israel
e-mail: michall@cc.huji.ac.il

U. Schmitz et al. (eds.), *Systems Biology of MicroRNAs in Cancer*, Advances in Experimental Medicine and Biology 1385,
https://doi.org/10.1007/978-3-031-08356-3_5

Abbreviations

CAGE	Cap-based expression analysis
ceRNA	Competing endogenous RNA
ChIP	Chromatin immunoprecipitation
circRNA	Circular RNA
CLASH	Cross linking, ligation and sequencing of hybrids
CLIP	Cross-linking immunoprecipitation
CNV	Copy number variation
DRV	Disease-related variation
FFL	Feed-forward loop
GEO	Gene expression omnibus
GO	Gene ontology
GWAS	Genome wide association study
HTP	High throughput
KEGG	Kyoto encyclopedia of genes and genomes
lncRNA	Long non-coding RNAs
LTP	Low throughput
MBS	miRNA-binding sites
miRNA	microRNA
ML	Machine learning
mRNA	Messenger RNA
MS	Mass spectrometry
MTI	miRNA-target interaction
RISC	RNA-induced silencing complex
RPM	Reads per million
Seq	Sequencing
smRNA	Small RNA
SNV	Single nucleotide variation
SVM	Support vector machine
TCGA	The Cancer Genome Atlas
TF	Transcription factor
TFBS	TF binding sites
TSS	Transcription start sites
UTR	Untranslated region

5.1 Human miRNA Regulation

Molecular View In multicellular organisms, microRNAs (miRNAs) play a role in driving cell differentiation, identity, and physiology (Wienholds and Plasterk 2005). A miRNA prototype is a single-stranded RNA molecule of approximately 22-nucleotide length that hybridizes to the 3′-UTR of its target transcript. In humans,

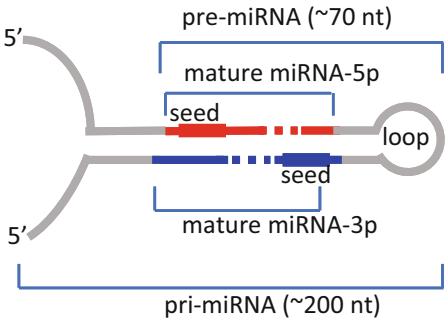


Fig. 5.1 A schematic view of stem-loop structure of the primary miRNA, pre-miRNA, and the mature miRNA products. The miRNA names are indicative of the source of the sequence from the stem-loop. For example, hsa-miR-142-5p and miR-142-3p are from the 5′ and 3′ arm, respectively. The dashed sign indicates non-perfect base-pairing in the stem-loop. The seed is 6–8 nt at the 5′ end of the mature miRNA. Often only one arm of the pre-miRNA is selected (guide strand) and loaded onto the RISC to form the miRISC

the ~1900 miRNA genes account for ~2600 mature miRNAs (Djuranovic et al. 2011). The RNA-induced silencing complex (RISC) protects the miRNA from degradation while stabilizing the miRNA–mRNA duplex. Consequently, the paired RISC–miRNA–mRNA shifts the protein translation efficiency by interfering with initiation, elongation, or termination steps. Occasionally, the miRNA–mRNA duplex also activates protein degradation processes. The bound transcript itself may undergo deadenylation, decapping, and further processing (e.g., degradation) (Cai et al. 2009).

The 5′ sequence of mature miRNAs includes the seed region (6–8 nt) that anchors the miRNA to the miRNA-binding sites (MBS) by a perfect base-pairing (Peterson et al. 2014) (Fig. 5.1). However, other sequence-based and structural features govern the actual binding specificity and efficiency. The fraction of the human transcriptome that is subjected to miRNA regulation is unknown. It is estimated that ~60% of the transcripts are regulated and have at least one conserved miRNA-binding site (MBS) at the 3′-UTR. Approximately 50% of all human protein-coding genes harbor alternative polyadenylation sites, resulting in transcripts with different 3′-UTR lengths (Müller et al. 2014). Naturally, alternative transcripts that differ in their 3′-UTR occur at different cell types. Consequently, the actual network of miRNAs and their targets is cell dependent. Whereas some mRNAs lack MBS, others may contain tens of them. From the miRNA perspective, while some miRNAs may pair with a limited number of targets, other miRNAs can pair with 100s of different mRNAs. Moreover, among the ~2600 reported human miRNAs, many have low expression levels and were only identified in specific cellular contexts by NGS (next-generation sequencing) data with an increased depth. Many of these candidate miRNAs lack experimental validation.

Cellular View For most cell systems, a detailed description of the transcriptome (i.e., mRNAs and miRNAs) allows determining each cell type and its origin (Gebert

and MacRae 2019). However, knowledge regarding the cell state and its physiology remains untraceable (Fu et al. 2013). Furthermore, regulating transcription factors (TFs) by miRNAs raises the need for assessing the direct and indirect effects within cells and tissues. In steady-state, miRNAs act as molecular agents for attenuating transcriptional noise. Upon changes in external conditions, miRNAs exploit cooperative and competitive modes that are difficult to model. For yielding an accurate model of cell homeostasis, evaluating the robustness of each cellular system to miRNA perturbations is needed (Mahlab-Aviv et al. 2019). Concretely, the molecular interactions of a miRNA with its targets can lead to abrupt changes in cell fate due to alterations in the levels of TFs, protein production, signaling cascades, and cell energetics (Alvarez-Garcia and Miska 2005).

Within living cells, sequestering of miRNAs by pairing with long noncoding RNAs (lncRNAs) leads to an apparent depletion in free miRNAs (sponge-like function). This implies an additional layer of complex regulation driven by the miRNA interactome (Militello et al. 2017). As the competition among miRNAs governs cell physiology, quantifying the amounts and the stoichiometry of miRNAs and mRNAs within cells is critical. The stochastic nature of miRNA-target interaction argues for using a probabilistic framework to describe the dynamics and the steady state of miRNAs and transcripts in living cells (Mahlab-Aviv et al. 2019). Capturing the bound miRNA–mRNA pairs yields a comprehensive and quantitative view of miRNA competition within living cells. Currently, most available miRNA tools fail to address the complexity of miRNA–mRNA pairing within cells. The contribution of miRNAs to the communication among neighboring cells was reported for neurons and glial cells (Morel et al. 2013). The generality of the miRNA-dependent signaling between cells awaits further studies. Merging cell studies (e.g., clinical tissues, cell lines, and single cells) with computational and experimental resources and tools is fundamental to empowering miRNA research.

5.2 The Scope and Organization of the Chapter

Plenty of resources and web tools were developed over the last 18 years (since 2003) for supporting miRNA research (Gomes et al. 2013). Studying miRNA regulation had been expanded along with the maturation of deep sequencing and diverse cross-linking immunoprecipitation (CLIP)-seq technologies. Many of the early developed tools aimed to predict miRNA–mRNA pairs. In reality, the many miRNA–mRNA prediction tools suffer from low consistency between them. Notably, results from computational prediction tools and experimental results show a high degree of inconsistency.

In recent years, numerous review articles have covered the collection of miRNA databases and tools (Shukla et al. 2017; Lukasik and Zielenkiewicz 2019; Akhtar et al. 2016; Shaker et al. 2020). Other publications concentrate on computational miRNA–mRNA prediction tools and their underlying algorithms (Mendes et al. 2009; Schmitz and Wolkenhauer 2013; Riffo-Campos et al. 2016; Monga and

Kumar 2019). A recent survey of the literature revisited ~100 review articles that covered ~1000 tools related to the broad field of miRNA (Chen et al. 2019). Unfortunately, many of the original tools and resources are discontinued or unstable. We focus on online tools and webservers and will not discuss stand-alone tools.

We aim to present a contemporary list of selected resources and online tools for studying miRNA regulation in health and disease. To allow an entry point to human-centric research, we will briefly mention tools in the context of the competition of miRNA and other noncoding RNAs (ncRNA) such as pseudogenes and circular RNAs (circRNAs). We will not discuss miRNA-related tools that focus on comparative genomics and evolution conservation (e.g., *miROrtho* (Gerlach et al. 2009) and *CoGemiR* (Maselli et al. 2008)).

The chapter starts by introducing a gateway to a human-centric collection of miRNA resources and online tools. We limit ourselves to those developed or updated in the last decade (from 2012) and are fully functional. We highlight tools that associate miRNAs with their targets according to computational and experimental approaches. We focus on tools that apply method integration, including major validated benchmarks. We then discuss resources that highlight dysregulation of miRNA in human diseases, specifically in cancer. In discussing the online tools, we consider the most updated version, as reported by primary publications. Finally, we will briefly review useful databases and online tools that are valuable in solving real-life experimental tasks regarding miRNA regulation in cellular systems.

5.3 Repositories for miRNA: Catalogs and Genome Browsers

MiRNAs are processed from hairpin-containing primary transcripts of ~200-nt (pri-miRNA) that are further processed to a ~70-nt stem-loop structure (pre-miRNA). These transcriptional events occur in the nucleus. All mature miRNAs undergo these biogenesis maturation steps. The pre-miRNA is then transported into the cytoplasm where a set of sequential cleavage events result in a functional miRNA that is bound to the RISC complex (Fig. 5.1). However, for many observed short RNAs that were identified from large-scale deep sequencing experiments, the definition of miRNAs is less definitive, and often indirect evidence from sequence conservation and independent experimental identification remains the sole support.

Several repositories for human miRNAs have been developed over the past 18 years. In the early days, the *microRNA Registry*, a branch of *Rfam* (Griffiths-Jones 2004; Kalvari et al. 2018) was compiled to facilitate the development of computational approaches for miRNA-target prediction. This registry was the basis for the current miRBase catalog (Kozomara et al. 2019). At present, miRNA notations and nomenclature are unified and adopted by the research community as presented in miRBase (Fromm et al. 2015).

5.3.1 miRNA Gene Catalogs

miRBase (release 22.1; 10/2018) is an exhaustive and inclusive miRNA catalog that aims to reach completeness. The miRNA collection was initially developed in 2006 and was regularly updated (Kozomara and Griffiths-Jones 2011). Presently, it includes over 1900 human miRNA genes and the notations for >2600 mature miRNA as observed from experimentally sequenced miRNAs. With the expansion of deep sequencing data, evidence from experiments is reported as normalized values (reads per million, RPM). Each miRNA is labeled on the pre-miRNA (stem-loop structure) and the mature processed version is depicted. A graphical viewer aligns the clustered sequences on the pre-miRNA and uses a unified nomenclature for the 3p and 5p arms (Fig. 5.1). Each miRNA in miRBase is associated with relevant publications and a detailed sequence of the precursor. Moreover, miRNAs are assigned to their families. The family members are miRNAs that derive from a common ancestor and have similar physiological functions (but are not necessarily conserved in sequence or structure) (Kamanu et al. 2013). Besides, miRBase provides a predicted secondary structure for miRNA hairpin loop precursors based on the RNAfold software. A confidence comment was added to allow the community to indicate the subjective confidence for the validity of a candidate miRNA. miRBase reports also on neighboring miRNAs by their chromosomal locations (i.e., miRNA clusters).

miRBase search engine allows extracting all cell or tissue-specific experiments. It is a useful entry point for miRNAs that were originally reported in *RNAcentral* (The RNAcentral Consortium et al. 2017). *miRBase Tracker* allows to keep track of historical and current miRNA annotations (Van Peer et al. 2014). miRBase FTP downloads allow the user to select data from any organism of choice (total 270).

MirGeneDB (Ver 2.0, 1/2020) is a robust platform for experimental results on small RNA (sRNA). While miRBase provides an exhaustive list of miRNA candidates, it suffers from a high level of false-positive entries. MirGeneDB aims to increase miRNA identification reliability by testing similarity among 45 metazoan species. The challenge for MirGeneDB is to provide an accurate assignment of miRNAs among expressed smRNA fragments derived from other genes (e.g., tRNAs, small nuclear (snRNAs), and small nucleolar RNAs (snoRNAs), piRNAs). The input starts with ~400 publicly available smRNA sequencing datasets extracted from *smRNAbench* (originally called *miRanalyzer*) (Aparicio-Puerta et al. 2019), and processed by miRTrace. A uniform annotation for each species relies on *MirMiner* that uses data derived from experiments of manipulated miRNA biogenesis genes (Fromm et al. 2020). Therefore, MirGeneDB also considers miRNA variants derived from noncanonical biogenesis. The current version compiled 556 annotated human miRNA genes that can be browsed, searched, and downloaded.

miRCarta (Ver 1.1 7/2018) is a database that features miRNA and precursor data from miRBase but complements the list from deep sequencing NGS from *miRMaster* and selected publications (Backes et al. 2018). The goal of miRCarta is to compile a consistent collection of novel miRNA candidates and augment the

information reported by miRBase. The database also includes an integrated information on predicted and experimentally validated targets extracted from *miRTarBase* (Huang et al. 2020), *microT-CDS* (Paraskevopoulou et al. 2013), and *TargetScan* (Agarwal et al. 2015). The functional links of miRCarta include the pathway dictionary of *miRPathDB* (Backes et al. 2016), and several miRNA-disease association databases (miR2Disease (Jiang et al. 2009), HMDD (Huang et al. 2019)). Besides, miRCarta provides a comprehensive collection of human miRNAs and miRNA candidates. It covers ~40 k miRNAs and precursors which are compressed to 2.9 k genomic clusters and 590 miRNA families.

5.3.2 *miRNAs in Genome Browsers*

An easily accessible entry point for the collection of miRNAs is supported by the major human genome browsers (e.g., UCSC and Ensembl). Figure 5.2 displays an overview of the main sources of primary data used in miRNA research. The primary data derived from NGS (next-generation sequencing) for genomics and transcriptomics (e.g., RNA-seq and smRNA-seq). The main genomic browsers provide the researcher with a comprehensive and up-to-date repository of genetic

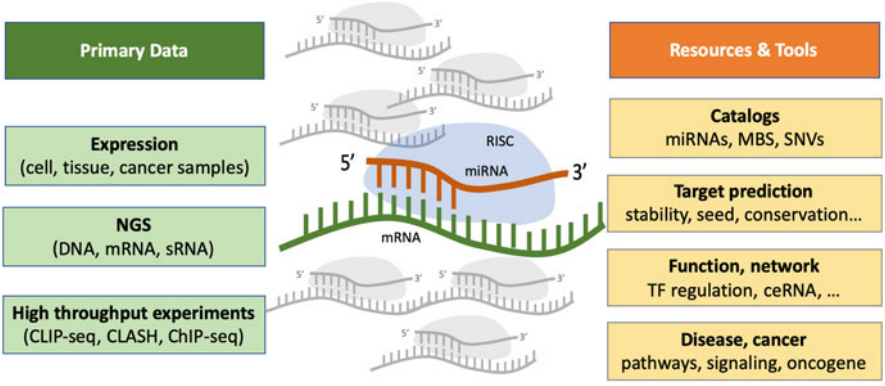


Fig. 5.2 An overview about sources of primary data for miRNA research combined with generic catalogs. The primary data is driven by recent advances in NGS (next generation sequencing) for genomics and transcriptomics (RNA-seq and smRNA-seq). NGS is applied to diverse biological samples. Specific methodologies for isolated miRNA and their targets include high-throughput (HTP) CLIP experiments (e.g., HITS-CLIP, iCLIP, and PAR-CLIP), CLASH based on miRNA-mRNA ligation protocol, ChIP-seq with TF and more. Rich experimental protocols of low throughput (LTP) include miRNA overexpression (OX), antimiR settings for downregulation of miRNA expression, mass spectrometry (MS), immunoprecipitation (IP) by RISC proteins, and numerous molecular manipulations with designed reporter (e.g., luciferase) for quantifying miRNA-dependent in vivo gene regulation (Thomson et al. 2011). Generic catalogs include the annotation of genes, alternative splicing (AS), Refseq transcripts, catalogs of cancerous cell lines (e.g., CCLE (Ghandi et al. 2019)), and human-related pathways (e.g., KEGG and BioCarta)

variants, transcripts, cross-species information, and a multidimensional rich data on cell regulation (Fig. 5.2).

The UCSC browser (5/2018) provides a collection of miRNAs as part of an annotation track for snoRNAs and miRNAs (Fujita et al. 2010). The data is matched to genome coordinates from miRBase, with only perfect matches (100% identity) are annotated. A list of miRNAs and genome coordinates are cross-referenced to miRBase. Together with the snoRNAs, it reports on 2230 gene name entries.

The Ensembl browser (Release 102, 11/2020) provides a collection of miRNAs that are annotated as part of the ncRNAs (Aken et al. 2016). Details are extracted from miRBase, with ~1900 entries for gene annotations. Each miRNA is considered by its functional arm of the pre-miRNA (3p or 5p; e.g., hsa-miR-155-3p, Fig. 5.1). When information is available, the relevant miRNA transcript is reported with evidence extracted from TarBase v8 (Karagkouni et al. 2018). The download of miRNA gene information of transcripts is supported by the Ensembl BioMart retrieval system.

5.4 Gateways for miRNAs: Integrative Platforms

The overwhelming number of tools and resources for miRNA analysis calls for easier access along with high-quality assessment. Elixir bio.tools registry (Ison et al. 2019) is such an entry point that provides a comprehensive registry of software and databases in all life science domains. Within the miRNA research domain, there are ~160 listed resources with comparable information and easy access. To ease the classification of the available web-based database, a meta-database was presented (*miRandb*) covering ~180 miRNA-centric databases (Aghaee-Bakhtiari et al. 2018). An entry point for human miRNAs is designed to specifically answer common miRNA-related research tasks. *HumiR* is an integrated website that compiled several human-centric databases and online tools for facilitating the selection of an appropriate tool (Solomon et al. 2020).

miRToolsGallery (9/2017) is a portal that provides an effective shortcut to main hubs for bioinformatics tools developed for miRNA research. The tool addresses the need for navigating among 100s of tools and the demand to meet the correct set of tools for any specific application. miRToolsGallery facilitates this process by curating ~950 miRNA analysis tools and resources. The portal provides a querying system for prioritizing the preferred tool for the needed application. The user can refine the search according to different criteria and requirements. Several features make this platform valuable as an entry point for miRNA research. One such feature allows flexible tagging of tools that belong to multiple categories. Moreover, it ranks results according to their popularity in citation and visibility (Chen et al. 2018a).

mirDIP 4.1 (1/2018) compiles a large number of computational miRNA-target prediction tools. It integrates >150 M human miRNA-target predictions across 30 different resources. Altogether, there are >48 M unique interactions, comprising ~2600 unique miRNAs and 28 k human genes. The database presents a

statistically-based integrative score for each miRNA-target interaction. Users can search for miRNA-target pairs according to the level of consistency between given resources, and apply several options as their desirable confidence score (Tokar et al. 2018).

Tools4miRs Ver 1.1 (3/2021) is a manually curated platform gathering all available tools that are miRNA related. Currently, there are 205 such tools (based on 170 methods), with the vast majority providing data on humans. The tools and database collections are categorized and further partitioned to more defined themes (e.g., isomiR identification and target functional analysis). Filtering by organisms, availability (e.g., online and downloading) facilitates the search for tools and databases that meet the user's needs. For example, the target prediction allows the user to define the target MBS positioning to 5′-UTR, coding, or 3′-UTR (Lukasik et al. 2016). The searched tools are presented in a simple or advanced mode. Tools4miRs also provide a meta-server for target prediction in which the user selects the designated methods to be included. It provides an option for reporting the miRNA-target prediction results via unification, intersection, or consensus method. A summary for each of the 205 tools is available along with a publication (Lukasik et al. 2016).

5.5 miRNA Gene Regulation: TFs and Cellular Context

The following set of resources are compiled from large-scale sequencing analysis with complementary information regarding gene expression and regulation. A unique feature unifying all these resources is the use of information from the cross-talk of miRNAs and cell-specific TFs. Many key resources benefit from cross-referenced complementary tools and algorithms and will briefly be mentioned. The *ChIPBase* database is a comprehensive annotation database from ChIP-Seq data mapping the transcriptional regulation of miRNAs (Yang et al. 2013). Other publicly available databases that address the problem of miRNA gene transcription regulation are *TFmiR* (Hamed et al. 2015), the tissue-specific miRNAs (*TSmiR* (Guo et al. 2014)), and more (Fig. 5.3).

DIANA-miRGen v4 (1/2021) is an updated version of experimentally supported functional relationships of miRNA-regulating genes (Alexiou et al. 2010). miRGen's goal is to provide an exhaustive resource for miRNA transcription start sites (TSS) extracted from a cap-based expression analysis (CAGE) of gene expression as reported by FANTOM (Abugessaisa et al. 2021). The TSS analysis covers most miRNA genes (1534 pre-miRNAs annotated in miRBase) across 135 different cellular contexts of diverse tissues, primary cells, and cell lines. Information on miRNA TSS is combined with the ENCODE ChIP-Seq results for TF binding sites (TFBS > 280) available from (Davis et al. 2018). miRGen provides detailed information on the genomic context of miRNAs, TF regulation (with multiple lines of experimental evidence), and cell-specific gene expression. It compiles a rich

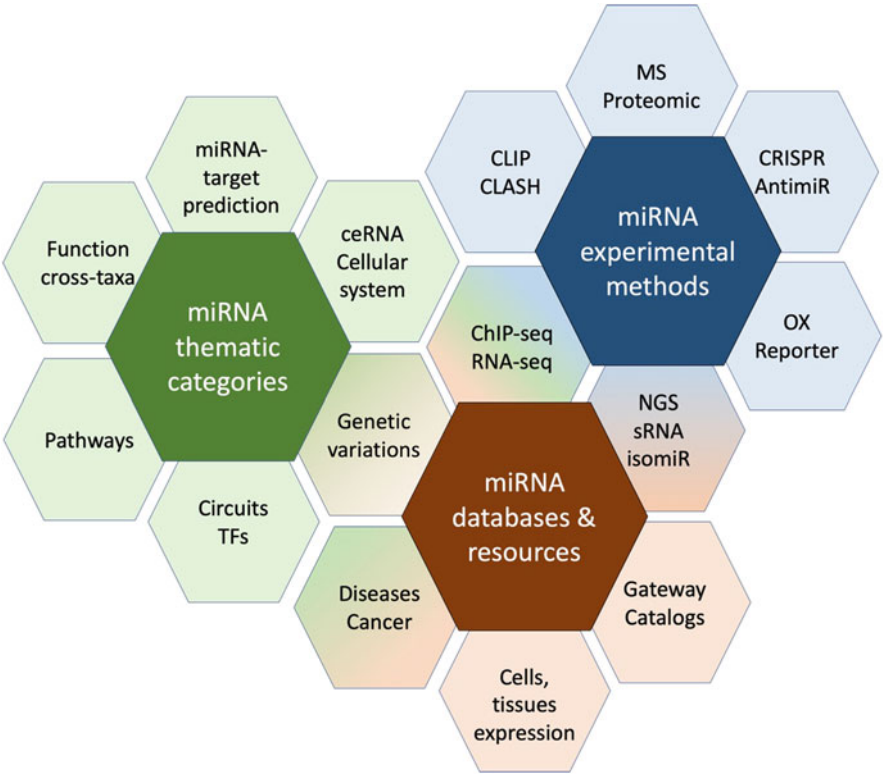


Fig. 5.3 An overview of main components in miRNA research: thematic categories, experimental methods, and the online tools, database, and resources. Experiments are based on miRNA-specific (e.g., CLASH) and generic methodologies (e.g., RNA-seq and MS proteomics). Cell perturbations for investigating miRNAs in cellular systems are based on overexpression (OX), downregulation (CRISPR, AntimiR), and designed reporters (e.g., luciferase) for assessing 3′-UTR regulation of target genes. Databases that were developed to facilitate miRNA research often rely on publicly available genomics, transcriptomics, and proteomics results from large-scale experiments (e.g., TCGA)

resource for miRNAs through cell-specific promoters and transcription regulation (Perdikopanis et al. 2021).

TransmiR v2.0 (1/2019) is a database that provides comprehensive information on TF-miRNA regulation based on surveying the literature and manual curation of >1300 publications. In addition, 1.7 M tissue-specific TF-miRNA regulations from ChIP-seq data were included. Querying the predicted TF-miRNA regulations in humans is based on information of the TF binding motifs. Additional capacities of TransmiR allow presenting the TF-miRNA interaction through a network or disease-centric views. Querying the system with a set of miRNAs allows the identification of TFs which most likely regulate this set of miRNAs. TransmiR provides a rich resource for investigating the regulation of miRNAs (Tong et al. 2019).

CircutesDB (1/2014) is a database of regulatory circuits composed of miRNAs and TFs. This resource integrates transcriptional and posttranscriptional regulatory networks. The basic notion is the existence of several circuit motifs such as feed-forward loops (FFLs) in which a TF regulates a miRNA, and together the expression level of a joint target is determined. A unique feature allows browsing among the catalog of regulatory motifs. Examples are intronic miRNA-mediated self-loops (iMSLs) (Friard et al. 2010).

5.6 miRNA-Target Prediction: Experiments and Validations

Most computational efforts and available tools for miRNA research address the pressing problem of target prediction. Namely, the task of accurately mapping miRNAs to their designated targets in a cellular context (Fig. 5.3). The many-to-many relationships make this problem challenging from a computational and experimental perspective. We will not elaborate on any of the algorithms behind target prediction tools that were thoroughly discussed (Schmitz and Wolkenhauer 2013; Riffo-Campos et al. 2016; Monga and Kumar 2019).

In a nutshell, the main features that are used by almost any of the tens of available miRNA-target prediction tools consider the degree of base-pairing with the seed region at the 5′ sequence with the mRNA (Fig. 5.1) (Biggar and Storey 2015). Additional features include the thermodynamic stability determined via the predicted minimum free energy of a putative miRNA-target duplex (Yue et al. 2009), and the estimated energy for removing the secondary structure of the target mRNA for exposing hidden MBS (i.e., MBS accessibility). Sequence conservation and genomic information across different branches of the taxonomy tree are utilized for removing false annotations, assigning reliability measures for the miRNA-target pairing (Peterson et al. 2014). Many more subtle features are considered for assessing the likelihood of the miRNA-target pairing. In some of the routinely used tools, a machine learning approach (e.g., support vector machine—SVM) was applied along the miRNA-prediction process. Most tools use scores to internally rank their predictions. Unfortunately, tool-specific scores are not easily generalized and rarely used in an integrative approach (Friedman et al. 2014). Recently, an application based on a semi-supervised ML method (*RPmirDIP*) confirmed the benefit of using prediction scores (Kyrollos et al. 2020).

5.6.1 miRNA-Target Prediction Tools and Resources

Selected tools for target prediction and validation, developed and updated since 2013, are listed in Table 5.1. We indicate the class of the algorithms and features that

Table 5.1 Selected miRNA-target prediction online tools and validation resources

Tool	Version	Update	Algorithm[a]	Taxa	Goal	References
miRGator	v3.0	2013	IA	1	Predict	Cho et al. (2013)
miRecords	v4	2013	Text	Many	Validate	Xiao et al. (2009)
miRTarCLIP	v1.0.1	2013	Exper, Seq	2	Predict	Chou et al. (2013)
mirMAP	V1.0	2013	IA	8	Predict	Vejnar et al. (2013)
PolymiRTS	v3.0	2014	IA, Seq	1	Predict variation	Bhattacharya et al. (2014)
miRror	v2.0	2014	IA	5	Predict	Friedman et al. (2014)
RNA22	v2.0	2015	Seq	Many	Predict	Loher and Rigoutsos (2012)
miTALOS	v2	2016	IA	2	Predict pathway	Preusse et al. (2016)
MiRTDL		2016	ML	1	Predict	Shuang et al. (2016)
TarBase	v8	2017	IA, Text	Many	Validate (Exper)	Karagkouni et al. (2018)
miRTar2GO		2017	Exper, Seq	1	Predict	Ahadi et al. (2017)
TargetScan	v7.2	2018	Seq	Many	Predict	Agarwal et al. (2015)
miRWalk	v3.0	2018	IA, Text	3	Predict	Dweep and Gretz (2015)
mirDIP	v4.1	2018	IA	1	Predict pathway	Tokar et al. (2018)
miRDB	v6.0	2019	ML	5	Predict	Chen and Wang (2020)
miRTarBase	2020	2020	Text	Many	Validate (Exper)	Huang et al. (2020)

[a]Algorithm used is based on an integrative approach (IA); Text mining/manual literature curation (Text); Machine learning, including deep neural networks (ML); Sequence features, matching seed, base-pairing complementarity, stability (Seq); Direct experimental data (Exper)

led to the development of the listed tools. Notably, many of these tools use machine learning (ML) that includes 100s of features and experimental results, allowing the algorithms to optimize the weights for each input data (Zou et al. 2015). As training by deep-learning efforts benefits from an increase in datasets, tools that rely on ML often give a superior prediction.

starBase v2.0 (1/2014) database reports on miRNA interactions as extracted from 108 CLIP-Seq experiments (PAR-CLIP, HITS-CLIP, iCLIP, CLASH). It covers 9 k miRNA–circRNA, 16 k miRNA–pseudogene, and 285 k protein–RNA regulatory relationships. starBase v2.0 provides a comprehensive pairing list for miRNA–mRNA and miRNA-lncRNA interactions based on CLIP-Seq data. A unique feature of starBase v2.0 is the miRFunction and ceRNAFunction servers that allow predicting the function of miRNAs and ncRNAs in the context of their regulatory networks and their coordinated function (Li et al. 2014).

miRDB v6.0 (1/2020) is an online database for miRNA target prediction and functional annotations (Chen and Wang 2020). The comprehensiveness of the

training data is evident in the RNA-seq dataset that consists of ~1.5 billion reads. Additionally, features extracted from CLASH (CLIP-ligation protocol) are considered as a validated set for the miRISC-MBS pairs. It covers the expression profiles of >1000 human cell lines and presents target prediction tailored for specific cell models. The underlying algorithm is MirTarget that was developed for analyzing NGS experiments. The predictions are based on a support vector machine (SVM) model. The prediction scores range from 0 to 100 where transcripts with scores of $\geq$50 are presented as predicted targets. For humans, miRDB covers ~2600 miRNA for 29 k genes leading to 1.6 M pairs. On average, there are ~600 gene targets per miRNA in humans. The miRDB is supplemented with a rich querying system with functional annotations and expression profiles from 100s of cell lines. All can be analyzed for GO annotation functional enrichment.

TargetScan v7.2 (3/2018) is designed to predict miRNA-targets by searching for the presence of conserved sequence features (e.g., seed and its extended variants) that match MBS in the 3′-UTR of mRNA transcripts. It combined a rich set of sequence-based features including the location of MBS relative to the stop codon. TargetScan also sorts miRNAs by their family relation according to the degree of taxonomical conservation (e.g., only mammals). The user can select to activate the analyses while choosing the level of miRNA family conservation (e.g., highly conserved in mammals, broadly conserved, or poorly conserved among mammals), including MBS with mismatches within the seed region. In mammals, the internal scoring reflects the predicted efficacy of targeting. The other scoring system shows the prediction ranking by the probability of conserved targeting (Agarwal et al. 2015).

miRTar2GO (4/2017) is a model trained on the accepted rules of miRNA–target pairing, including experimentally validated interactions from CLIP-seq data. miRTar2GO allows the prediction of cell-type-specific targets. The model provides biological insights through GO enrichment of miRNA-targets (Ahadi et al. 2017). miRTar2GO ranks the interactions predicted for a miRNA based on its distance to the verified interactions of that miRNA. A unique feature is the option to activate the model as highly specific or highly sensitive (Ahadi et al. 2017).

miRGate (4/2015) is a curated database of miRNA–mRNA targets that compare several established miRNA-related experimental databases and integrate major miRNA-target prediction tools (*microTar*, *RNAHybrid*, *miRanda*, *TargetScan*) (Andres-Leon et al. 2015).

miRror-Suite (6/2014) is an integrative set of tools dealing with miRNA regulation in the context of living cells. Specifically, it allows a query on a set of differentially expressed miRNAs that list the most likely targets that are affected by such a set. It allows the user to reanalyze the data by redefining the statistical significance thresholds. It is based on miRror v2.0 that also provides the opposite view. Namely, from a set of differentially expressed mRNAs as input, find the most likely miRNAs set that plays a role in the regulation. The miRror-Suite *miRtegrate* algorithm designates statistical criteria that were uniformly applied to a dozen miRNA-target prediction databases. The user can refine the analysis by selecting

the desired tissues, cell lines, differential expression, and internal predicting scores (Friedman et al. 2014).

miRGator 3.0. (1/2013) serves as a miRNA portal that relies on NGS data for miRNA diversity, expression, and target relationships. It is based on 73 NGS datasets from major gene expression resources (GEO, SRA, and TCGA) that include 2.5B aligned reads. The database provides expression data by anatomical description and assigned the miRNA data to 38 human diseases with summary statistics (Cho et al. 2013). A unique feature is the availability of tools to facilitate the exploration of massive raw NGS reads for finding miRNAs. By using the *miRDeep2* algorithm novel miRNAs, isomiRs, and edited miRNA versions are sought. Moreover, the portal allows comparing gene sets from different studies and extracting biological insights by functional enrichment scheme and gene set analysis.

The utilization of publicly available NGS data such as transcriptomic data for successful use by the miRNA community is challenging. Many databases (e.g., microRNA.org, *deepBase*, and *miRBase*) quantify the results from smRNA-seq data for presenting normalized expression values. Many computational and bioinformatics tools combined HTP experimental data with normalized and processed raw data. Still, a set of tools were developed to assist researchers in using miRNA-specific NGS-based pipelines.

miRMine (5/2017) compiled ~350 miRNA-seq datasets from NCBI-SRA. miRMine reported on ~2500 mature RNAs and their RPM (reads per million) expression level for 16 human tissues including body fluids and 24 cell lines (Panwar et al. 2017).

smRNAtoolbox (5/2015) provides a collection of useful tools for NGS experiment analyses (*smRNAbench*), complemented with several miRNA analysis tools. While it is not a miRNA dedicated platform, it contains seven independent tools that create a workflow for miRNA analyses. The tools are designed to meet a realistic flow for NGS miRNA-seq experiments. Integration of tools allows the user to benefit from a set of established smRNA bioinformatic tools for read mapping, expression profiles, differential expression, genome browser visualization, enrichment of functional annotations, pathway viewer, and cross-species miRNA target prediction (Rueda et al. 2015).

5.6.2 miRNA-Target Prediction Validation Databases

Results from experimental CLIP-seq studies, CLASH, and the advances in NGS led to a wave of datasets that are the basis for updating many miRNA-target prediction tools. Such an effect led to high-quality miRNA-target validated resources as benchmarks in the miRNA field.

DIANA-TarBase v8 (1/2018) database reports on experimentally supported miRNA targets. This resource is considered a benchmark for several prediction methods. The current TarBase reports on ~670 k unique miRNA-target pairs. The

database compiles information from a large set of experimental methodologies, conditions, and cellular contexts, covering about 600 cell types and tissues. The database provides an interactive querying system and rich filtering options in addition to the browsing capacity. Retrieval of miRNA-target pairs is activated according to a combined selection of species, supporting methodology, and the selected cell type. TarBase v8 presents a ranking system that is based on the empirical reliability of the method used as evidence (Karagkouni et al. 2018).

miRTarBase 2020 (1/2020) is a comprehensive resource of experimentally validated miRNA-target interactions (denoted MTIs). miRTarBase is a rich experimentally validated MTI database with comprehensively annotated information (Huang et al. 2020). The database covers >380 k validated MTIs for humans. Such MTIs are based on ~2600 miRNAs and >15 k targets with supporting evidence from 7.2 k manually curated publications. A scoring system based on text mining ranks any miRNA-target interaction pair. Evidence from direct assays (e.g., Western blot, qPCR, and reporter gene) are scored higher than those from large-scale methodologies (e.g., CLIP-Seq). Also, a large number of databases were integrated to provide rich information on the number of MTIs within a regulatory network (based on KEGG pathways). The database also provides the current list of validated miRNA-targets according to CLIP-seq technology (Huang et al. 2020).

miRecords (4/2013) consists of experimentally validated miRNA-targets as revealed from literature curation. In addition, it provides a synthesis of many target prediction programs (e.g., *DIANA-microT*, *miTarget*, *PITA*, and *TargetScan*). miRecords hosts over 2700 records of miRNA-target pairs, with information from direct testing of interaction. It covers about 650 miRNAs and 1900 target genes (Xiao et al. 2009).

5.7 miRNA-Target Databases: Networks and Pathways

On the basis of the predicted and experimentally validated miRNA-target interactions, several databases were developed to address complex regulatory networks in the context of cellular pathways. Assignment of miRNAs to pathways according to individual prediction tools (Table 5.1) suffers from an uncontrolled number of false-positive predictions and poor level of agreement. Relying on the consistency of miRNA-target prediction tools and predetermined topology of human pathways showed that miRNAs are crucial in most pathways from KEGG and the pathway interaction database (PID) (Naamati et al. 2012).

mirDIP v4.1 (1/2018) provides nearly 152 M miRNA-target predictions collected from 30 different resources. The underlying algorithm *NAViGaTOR* (Shirdel et al. 2011) produces miRNA-target networks from the literature and pathways databases (e.g., KEGG and Reactome). The signaling pathway networks that are signified by miRNA involvement are listed and scored (Tokar et al. 2018).

miRWalk v3.0 (10/2018) is a platform providing an intuitive interface that generates predicted and validated MBS. miRWalk uses a random-forest approach implemented in the *TarPmiR* algorithm to search for MBS across the entire transcript length (i.e., MBS is not limited to the 3′-UTR). The current version of miRWalk stores predicted data including experimentally verified miRNA-target interactions (Sticht et al. 2018). The human version covers the entire miRNA set (2656 miRNAs according to miRBase Ver 22). The pairing is with regard to the ~20 k RefSeq coding genes and 42 k transcripts. miRWalk provides a cross-reference to other major miRNA-target predictions (e.g., *TargetScan* and *miRDB)*. The pairwise miRNA–mRNA 100 M reported interactions include the ~0.9 M validated pairs from *miRTarBase database* (Sticht et al. 2018).

miRPathDB v2.0 (1/2020) is a dictionary of miRNAs and pathways. It covers an exhaustive collection of candidate miRNAs from miRBase v.22.1 and miRCarta (v.1.1), 28 k human targets and ~17 k pathways for Homo sapiens. It uses the validated MTIs from *miRTarBase* and activates target prediction by using *TargetScan* (v.7.1) and *miRanda*. A querying system allows determining the maximal number of miRNAs to be presented based on a reference pathway (e.g., from KEGG, Recatome, and WikiPathways). In addition, it provides new functionality by allowing users to determine a threshold for the reliability of the results. The miRPathDB presents similarity maps for miRNAs and pathways by the statistical significance of overlapping in their targets and pathways. The visualization tools and the downstream analysis are designed to determine a minimal set of candidate regulators that are sufficient to target a gene list (Kehl et al. 2020).

5.8 miRNA Sponging: ceRNA and lncRNA Interactions

An indirect regulatory level of miRNA function is formulated by the concept of competing endogenous RNAs (ceRNAs). In cells, miRNAs may be sequestered by RNA molecules that contain MBS but are not genuine mRNA targets. These RNAs act as miRNA sponges and are often pseudogenes, long noncoding RNAs (lncRNA), or circRNAs. Cell physiological and pathological processes are often regulated by ceRNAs. To fully appreciate the in vivo steady-state in cells, the quantitative aspects of miRNAs in the cellular context and the subtleties of cellular states and molecular events such as miRNA partition between nucleus and cytoplasm, exosome signaling, and miRNA editing may impact the in vivo miRNA-target interaction landscape.

miRSponge (9/2015) is a manually curated database that provides experimentally supported resources for miRNA sponge interactions. miRSponge reports on 298 miRNA–ceRNA interactions in humans that occur between miRNA, pseudogenes, lncRNAs, circRNAs, and human viruses. The database covers 11 species with ~600 miRNA–ceRNA interactions that are supported by ~1200 published articles. miRSponge is a webtool with browsing, retrieval, and downloading

capacities. A unique feature is a submission page that allows researchers to enrich the resource by adding validated miRNA sponge data (Wang et al. 2015).

DIANA-LncBase v.2 (1/2016) is a database of experimentally supported and in silico predicted MBS in lncRNAs. DIANA-LncBase is an extensive collection of miRNA–lncRNA interactions with ~70 k experimentally supported interactions derived from publications and the analysis of ~150 AGO CLIP-Seq libraries. In addition to the experimentally validated set, DIANA-LncBase lists in silico predictions based on the *DIANA-microT* algorithm. A unique feature is the association of the prediction results with information regarding 66 different cell types from 36 tissues. The database includes an exhaustive analysis of six billion RNA-Seq reads for obtaining accurate cell-specific lncRNA expression information (Paraskevopoulou et al. 2016).

5.9 Genomic miRNA Databases: Variations and isomiRs

The following resources address the impact of human genome variations on miRNA regulation via changes in the identity and specificity of MBSs and miRNAs. Also, the immense NGS repository becomes fundamental for identifying isomiRs and other miRNA candidates (Glogovitis et al. 2020).

PolymiRTS v3.0 (1/2014) is a platform for analyzing the impact of genetic polymorphisms in miRNA seed regions and miRNA target sites for humans and mice. The resource provides a comprehensive list of naturally occurring genetic variations in seed regions of miRNAs and the MBS on targets. SNVs and indels in miRNAs and their MBS have the potential to alter miRNA-mediated regulation. This database is a useful resource for genotype and phenotype association studies. The data is based on CLASH experimental results of miRNA–mRNA interactions. Unique features include the use of polymorphic sites of *TargetScan* scores. Interpretation for the SNVs occurring at the 3′-UTR of the human transcripts are presented by searching the downstream effects on gene expression and pathways in the context of genome wide association studies (GWAS) for human traits and diseases (Bhattacharya et al. 2014).

miRdSNP (1/2012) aims to present the impact of SNVs with regard to diseases. SNVs could lead to gene dysregulation by modifying the efficiency of miRNA binding to the 3′-UTR of the target. miRdSNP is based on a manually curated literature survey with ~800 disease associations SNVs and ~200 disease types, an extended list of experimentally validated miRNA–mRNAs pairs, and sites predicted by key predicting tools (e.g., *TargetScan*). The tool allows searching for the distance of the identified MBS and SNVs associated with human diseases. It also provides a viewer through the UCSC Genome browser (Bruno et al. 2012).

MIRPIPE (8/2016) is a pipeline for the quantification of miRNA based on smRNA sequencing reads. MIRPIPE allows an automatic trimming of sequencing adapters from raw RNA-seq reads and clustering of isomiRs. MIRPIPE does not rely

on the generic reference genome to identify miRNAs. A unique feature is its flexible model for miRNA quantification. Any uploaded database can be considered a reference for the homology search engine (Kuenne et al. 2014).

SomamiR v 2.0 (1.2016) is a database of cancer somatic mutations that potentially affect miRNAs and their targets. It addresses the impact of genetic alterations on ceRNA, including their effect on circRNAs and lncRNAs. SomamiR provides an integrated platform for functional analysis of somatic mutations. To this end, *miR2GO* is used to analyze the functional consequences of somatic mutations in the seed region of miRNAs. Besides, experimental data (CLASH, CLIP-seq) are analyzed given the somatic mutations. The database highlights mutations in miRNAs and their target sites that change cancer risk as reported in GWAS and various experimental evidence (Bhattacharya and Cui 2016).

Enriching the miRNA variant landscape from external data collections led to the development of dedicated pipelines. An example is *miRVar* which is based on *LOVD v.2.0* (Build 22) (Bhartiya et al. 2011). A machine learning approach using SVM predicts the processing sites of pre-miRNA and the guide strand selection (Fig. 5.1). The possible effect of variations in miRNAs was assessed based on the expected penetrance in the human population.

5.10 miRNA Dysregulation: Diseases, Cancer, and Signaling

The regulation by miRNAs on their intended target only represents a snapshot of a dynamic circuit (Re et al. 2017). miRNAs take active parts in many pathologies and altered signaling pathways. The regulatory networks are produced from small network motifs that are recurrent in nature. While motifs that involve TFs were studied extensively, those including lncRNAs or epigenetic regulators (Sato et al. 2011) introduce overlooked dimensions to the role of miRNAs in cell physiology and pathology.

Regulation of gene expression is the key to maintaining homeostatic processes. Several databases aim to cope with miRNA-target interactions upon changing conditions (e.g., *CSmiRTar* (Wu et al. 2017). In this view, many events that involve cell pathological conditions are reflected by a shift in miRNA action in cells (e.g., *miRwayDB* (Das et al. 2018)). The dysregulation of miRNAs has been associated with many diseases such as type 2 diabetes (T2D) and cardiovascular diseases. The impact of miRNAs in other conditions such as arthritic diseases, Alzheimer, and several mental conditions becomes evident. The discovery of miRNA signaling by exosomes is a novel aspect of cell–cell regulation and is an attractive source of biomarkers.

5.10.1 Disease-Related miRNA Databases

Studying disease-related miRNAs is beneficial to understand disease mechanisms at the miRNA level. However, most current methodologies in miRNA research (Fig. 5.2) are limited to in vitro binding assays and cellular manipulation (Fig. 5.3). Several studies have developed networks of miRNAs and diseases (Gu et al. 2016). The validity of such networks was analyzed (Chen et al. 2018b), and proved to be useful for miRNA-disease relation predictions (You et al. 2017; Chen et al. 2018c).

HMDD v3.2 (1/2019) (Human microRNA Disease Database) is a database that curated experiment-supported evidence for miRNA-disease associations. The list of evidence (with 20 evidence codes) includes genetics, epigenetics, circRNAs, and miRNA-target interactions. HMDD bridges between observations from numerous experiments and disease-associated miRNAs. HMDD also covers GWAS results and copy number variations (CNV) leading to gain and loss of genomic miRNAs (Huang et al. 2019). The current version covers a manually collected list of 35.5 k miRNA-disease associations involving 1200 miRNAs and about 900 human diseases. The findings are supported by >19 k publications. The disease annotation is linked to ICD-10 that is the unified index used by the medical community. A connection to major disease ontology terms (e.g., *DOID*, *MESH*, *OMIM*, and *HPO*) is also provided (Huang et al. 2019).

miRNASNP-v3 (1/2020) is a rich resource that combines data on genetic variations in miRNAs and MBS with disease-related variations (DRVs). miRNASNP is used to determine the possible effect of SNVs on miRNA interactions. The resource analyzes >7 M germline and somatic SNVs and 0.5 M disease-related variations with respect to ~2600 mature miRNAs and > 18 k 3′ UTRs of human genes. miRNASNP compiled the set of SNVs from clinical samples (*ClinVar* and *COSMIC*) and population variation catalogs (dbSNP, *GWAS Catalog*). It provides a functional enrichment analysis of miRNA target gain/loss caused by SNPs/DRVs. Correlations between drug sensitivity and miRNA expression level are presented, with a special focus on potential targets in cancers (Liu et al. 2021).

miRandola (9/2017) is a database of extracellular ncRNAs that are attractive as noninvasive biomarkers from body fluids. miRandola collected data from 314 articles that reported on ~1000 miRNAs and other ncRNAs. The website provides a browsing capacity, name convertor, and details tabular information on the experiments and the nature of the carrier of miRNA (e.g., exosome and microparticle) (Russo et al. 2018).

5.10.2 Cancer-Related miRNA Databases

As miRNAs govern cell identity and physiology in many tissues, alterations in miRNAs signify all cancer types. Human cancer databases such as The Cancer Genome Atlas (TCGA) provide a rich resource for the expression levels of miRNAs and mRNAs for over 14 k cancer samples. Other collections include the OncomiR cancer database (e.g., (Sarver et al. 2018)). The expression levels of oncogenic miRNAs (oncomiRs) and those that act as tumor suppressors make them attractive sites for manipulation and a lead for cancer therapeutic methods.

dbDEMC 2.0 (1/2017) is a cancer-specific resource for storing and displaying differentially expressed miRNAs in human cancers. It uses a simple text search for human cancers from the GEO gene expression data collection. The latest version of dbDEMC contains ~2200 differentially expressed miRNAs identified for 36 cancer types (73 subtypes) from 436 experiments. From large-scale analyses of cancer samples (based on ~150 publications), a collection of 49 k miRNA–cancer associations is provided (Yang et al. 2017). For example, based on TCGA, a list of miRNAs in colon cancer (total 2100) is split into those induced or suppressed relative to the healthy tissue. A unique feature is a meta-profiling representation that allows the user to provide an input set of miRNAs and retrieves as an output their differential expression trend by a heatmap according to broad clinical characteristics (e.g., metastasis, high and low grades) (Yang et al. 2017).

miRCancer (1/2013) provides a comprehensive collection of miRNA expression profiles in various human cancers that are automatically extracted from published literature using text-mining approaches. It utilizes rule-based techniques for mining key sentences regarding the expression trend in cancer and control cells. Manual revision is applied after auto-extraction to improve precision. miRCancer reports on 236 miRNAs and 79 human cancer types from 26 k publications. A unique feature is the constant updating of the information by analyzing the literature in PubMed (Xie et al. 2013).

miRNACancerMAP (9/2018) is a user-friendly web server with integrated data sources and a computational workflow for exhaustive searching of miRNA-cancer information. Specifically, one can ask for the common miRNA-gene regulation networks in multiple cancers using context-dependent expression evidence. The resource allows identifying the sponge regulations by lncRNAs in a clinical setting. The interactive interface allows merging of public data (e.g., *TCGA* and *GEO*) with user results such as cancer-derived miRNA–mRNA expression data. Therefore, for the known pathways (e.g., *KEGG* and *Reactome*) the impact of miRNA dysregulation on cancer is determined. It allows highlighting miRNAs acting as cancer drivers and tumor suppressors by the cancer hallmark database. A unique feature is the possibility to analyze the user miRNA data by providing interactive visualization tools, and activating multiple miRNA algorithms (Tong et al. 2018).

OncomiR (2/2018) is a user-friendly resource for exploring miRNA dysregulation in cancer. OncomiR covers ~1200 mature miRNA and 30 k mRNA transcripts from

~10 k patients spanning 30 cancer types, along with statistical analysis. OncomiR consists of a database and a dynamic web server. Key functions of OncomiR are the identification of cancer-relevant miRNAs and de novo analysis based on miRNA expression. The unique functionality of OncomiR is in providing the most significant miRNAs for any specific cancer type. Moreover, it allows listing potential miRNAs for a survival signature with Kaplan–Meier (KM) survival curve representation available for a given cancer type (Wong et al. 2018).

5.11 Summary and Future Perspectives

The field of experimental and computational miRNA research has been gradually evolving over the last 18 years (Fig. 5.2). To this end, hundreds of stand-alone, online tools (e.g., *multiMiR* package) (Ru et al. 2014), algorithms, and databases have been developed for miRNA research. The main task was to provide simple rules for the miRNA regulation in living cells and at the organism level. Unfortunately, the degree of inconsistency remains high among the many miRNA-target prediction tools (Riffo-Campos et al. 2016). Therefore, selecting suitable databases and tools for researchers became increasingly challenging. In this chapter, we briefly discuss tools and databases for assisting miRNA-focused research according to major categories (Fig. 5.3). Notably, the improved HTP technologies such as deep sequencing led to an increase in the number of miRNA candidates, with many of them still awaiting experimental validation. It became clear that in living cells, examining miRNA profiles is not limited to simple miRNA-target pairing rules. Instead, an integration of different regulation layers (TFs, epigenetics, translation, and lncRNAs) is essential. Currently, tools for quantifying key players (i.e., miRNAs, TFs, mRNAs, and proteins) and their dynamics in living cells are missing (Mahlab-Aviv et al. 2019). Such measurements are essential for evaluating the degree of competition and cooperativity among miRNAs in cellular systems (Balaga et al. 2012). The use of medical informatics to determine genetic variations and their impact on diseases allowed to bridge between miRNA research and research in human health (Fig. 5.4). Refining experimental methods for miRNAs and collecting accurate data in databases and well-maintained and undated online tools will continue to advance the field. Specifically, designing solid benchmarks for comparing the tools' performance is a pressing need. The current knowledge on miRNA regulation in health and disease will benefit from modern statistical methods (e.g., deep learning) and further development of integrative approaches.

Fig. 5.4 A comparative table for selected set of disease-oriented online tools and databases. Each of the tools is indicated by its focus (any/specific disease or cancer), the use of large-scale experimental results, the capacity for downloading the pre-processed data and whether the resource is restricted to human or covers other species (marks as minus and plus sign, respectively). The year of the last major update is noted according to the primary publication

References

Abugessaisa I, Ramilowski JA, Lizio M, Severin J, Hasegawa A, Harshbarger J, Kondo A, Noguchi S, Yip CW, Ooi JLC (2021) FANTOM enters 20th year: expansion of transcriptomic atlases and functional annotation of non-coding RNAs. Nucleic Acids Res 49:D892–D898

Agarwal V, Bell GW, Nam JW, Bartel DP (2015) Predicting effective microRNA target sites in mammalian mRNAs. elife 4

Aghaee-Bakhtiari SH, Arefian E, Lau P (2018) miRandb: a resource of online services for miRNA research. Brief Bioinform 19:254–262

Ahadi A, Sablok G, Hutvagner G (2017) miRTar2GO: a novel rule-based model learning method for cell line specific microRNA target prediction that integrates Ago2 CLIP-Seq and validated microRNA–target interaction data. Nucleic Acids Res 45:e42–e42

Aken BL, Ayling S, Barrell D, Clarke L, Curwen V, Fairley S, Fernandez Banet J, Billis K, García Girón C, Hourlier T (2016) The Ensembl gene annotation system. Database 2016

Akhtar MM, Micolucci L, Islam MS, Olivieri F, Procopio AD (2016) Bioinformatic tools for microRNA dissection. Nucleic Acids Res 44:24–44

Alexiou P, Vergoulis T, Gleditzsch M, Prekas G, Dalamagas T, Megraw M, Grosse I, Sellis T, Hatzigeorgiou AG (2010) miRGen 2.0: a database of microRNA genomic information and regulation. Nucleic Acids Res 38:D137–D141

Alvarez-Garcia I, Miska EA (2005) MicroRNA functions in animal development and human disease. Development 132:4653–4662

Andres-Leon E, Gonzalez Pena D, Gomez-Lopez G, Pisano DG (2015) miRGate: a curated database of human, mouse and rat miRNA-mRNA targets. Database (Oxford) bav035

Aparicio-Puerta E, Lebron R, Rueda A, Gomez-Martin C, Giannoukakos S, Jaspez D, Medina JM, Zubkovic A, Jurak I, Fromm B, Marchal JA, Oliver J, Hackenberg M (2019) smRNAbench and smRNAtoolbox 2019: intuitive fast small RNA profiling and differential expression. Nucleic Acids Res 47:W530–W535

Backes C, Kehl T, Stöckel D, Fehlmann T, Schneider L, Meese E, Lenhof H-P, Keller A (2016) miRPathDB: a new dictionary on microRNAs and target pathways. Nucleic Acids Res gkw926

Backes C, Fehlmann T, Kern F, Kehl T, Lenhof H-P, Meese E, Keller A (2018) miRCarta: a central repository for collecting miRNA candidates. Nucleic Acids Res 46:D160–D167

Balaga O, Friedman Y, Linial M (2012) Toward a combinatorial nature of microRNA regulation in human cells. Nucleic Acids Res 40:9404–9416

Bhartiya D, Laddha SV, Mukhopadhyay A, Scaria V (2011) miRvar: a comprehensive database for genomic variations in microRNAs. Hum Mutat 32:E2226–E2245

Bhattacharya A, Cui Y (2016) SomamiR 2.0: a database of cancer somatic mutations altering microRNA–ceRNA interactions. Nucleic Acids Res 44:D1005–D1010

Bhattacharya A, Ziebarth JD, Cui Y (2014) PolymiRTS database 3.0: linking polymorphisms in microRNAs and their target sites with human diseases and biological pathways. Nucleic Acids Res 42:D86–D91

Biggar KK, Storey KB (2015) Insight into post-transcriptional gene regulation: stress-responsive microRNAs and their role in the environmental stress survival of tolerant animals. J Exp Biol 218:1281–1289

Bruno AE, Li L, Kalabus JL, Pan Y, Yu A, Hu Z (2012) miRdSNP: a database of disease-associated SNPs and microRNA target sites on 3′UTRs of human genes. BMC Genomics 13:1–7

Cai Y, Yu X, Hu S, Yu J (2009) A brief review on the mechanisms of miRNA regulation. Genomics Proteomics Bioinformatics 7:147–154

Chen Y, Wang X (2020) miRDB: an online database for prediction of functional microRNA targets. Nucleic Acids Res 48:D127–D131

Chen L, Heikkinen L, Wang C, Yang Y, Knott KE, Wong G (2018a) miRToolsGallery: a tag-based and rankable microRNA bioinformatics resources database portal. Database (Oxford)

Chen X, Wang C-C, Yin J, You Z-H (2018b) Novel human miRNA-disease association inference based on random forest. Molecular Therapy-Nucleic Acids 13:568–579

Chen X, Yin J, Qu J, Huang L (2018c) MDHGI: matrix decomposition and heterogeneous graph inference for miRNA-disease association prediction. PLoS Comput Biol 14:e1006418
Chen L, Heikkinen L, Wang C, Yang Y, Sun H, Wong G (2019) Trends in the development of miRNA bioinformatics tools. Brief Bioinform 20:1836–1852
Cho S, Jang I, Jun Y, Yoon S, Ko M, Kwon Y, Choi I, Chang H, Ryu D, Lee B, Kim VN, Kim W, Lee S (2013) MiRGator v3.0: a microRNA portal for deep sequencing, expression profiling and mRNA targeting. Nucleic Acids Res 41:D252–D257
Chou CH, Lin FM, Chou MT, Hsu SD, Chang TH, Weng SL, Shrestha S, Hsiao CC, Hung JH, Huang HD (2013) A computational approach for identifying microRNA-target interactions using high-throughput CLIP and PAR-CLIP sequencing. BMC Genomics 14(Suppl 1):S2
Das SS, Saha P, Chakravorty N (2018) miRwayDB: a database for experimentally validated microRNA-pathway associations in pathophysiological conditions. Database (Oxford)
Davis CA, Hitz BC, Sloan CA, Chan ET, Davidson JM, Gabdank I, Hilton JA, Jain K, Baymuradov UK, Narayanan AK (2018) The encyclopedia of DNA elements (ENCODE): data portal update. Nucleic Acids Res 46:D794–D801
Djuranovic S, Nahvi A, Green R (2011) A parsimonious model for gene regulation by miRNAs. Science 331:550–553
Dweep H, Gretz N (2015) miRWalk2.0: a comprehensive atlas of microRNA-target interactions. Nat Methods 12:697
Friard O, Re A, Taverna D, De Bortoli M, Cora D (2010) CircuitsDB: a database of mixed microRNA/transcription factor feed-forward regulatory circuits in human and mouse. BMC Bioinformatics 11:435
Friedman Y, Karsenty S, Linial M (2014) miRror-suite: decoding coordinated regulation by microRNAs. Database (Oxford)
Fromm B, Billipp T, Peck LE, Johansen M, Tarver JE, King BL, Newcomb JM, Sempere LF, Flatmark K, Hovig E, Peterson KJ (2015) A uniform system for the annotation of vertebrate microRNA genes and the evolution of the human microRNAome. Annu Rev Genet 49:213–242
Fromm B, Domanska D, Hoye E, Ovchinnikov V, Kang W, Aparicio-Puerta E, Johansen M, Flatmark K, Mathelier A, Hovig E, Hackenberg M, Friedlander MR, Peterson KJ (2020) MirGeneDB 2.0: the metazoan microRNA complement. Nucleic Acids Res 48:D1172
Fu G, Brkić J, Hayder H, Peng C (2013) MicroRNAs in human placental development and pregnancy complications. Int J Mol Sci 14:5519–5544
Fujita PA, Rhead B, Zweig AS, Hinrichs AS, Karolchik D, Cline MS, Goldman M, Barber GP, Clawson H, Coelho A (2010) The UCSC genome browser database: update 2011. Nucleic Acids Res 39:D876–D882
Gebert LF, MacRae IJ (2019) Regulation of microRNA function in animals. Nat Rev Mol Cell Biol 20:21–37
Gerlach D, Kriventseva EV, Rahman N, Vejnar CE, Zdobnov EM (2009) miROrtho: computational survey of microRNA genes. Nucleic Acids Res 37:D111–D117
Ghandi M, Huang FW, Jané-Valbuena J, Kryukov GV, Lo CC, McDonald ER, Barretina J, Gelfand ET, Bielski CM, Li H (2019) Next-generation characterization of the cancer cell line encyclopedia. Nature 569:503–508
Glogovitis I, Yahubyan G, Wurdinger T, Koppers-Lalic D, Baev V (2020) isomiRs-hidden soldiers in the miRNA regulatory Army, and how to find them? Biomol Ther 11
Gomes CPDC, Cho J-H, Hood LE, Franco OL, Pereira RWD, Wang K (2013) A review of computational tools in microRNA discovery. Front Genet 4:81
Griffiths-Jones S (2004) The microRNA registry. Nucleic Acids Res 32:D109–D111
Gu C, Liao B, Li X, Li K (2016) Network consistency projection for human miRNA-disease associations inference. Sci Rep 6:1–10
Guo Z, Maki M, Ding R, Yang Y, Xiong L (2014) Genome-wide survey of tissue-specific microRNA and transcription factor regulatory networks in 12 tissues. Sci Rep 4:1–9

Hamed M, Spaniol C, Nazarieh M, Helms V (2015) TFmiR: a web server for constructing and analyzing disease-specific transcription factor and miRNA co-regulatory networks. Nucleic Acids Res 43:W283–W288

Huang Z, Shi J, Gao Y, Cui C, Zhang S, Li J, Zhou Y, Cui Q (2019) HMDD v3. 0: a database for experimentally supported human microRNA–disease associations. Nucleic Acids Res 47: D1013–D1017

Huang HY, Lin YC, Li J, Huang KY, Shrestha S, Hong HC, Tang Y, Chen YG, Jin CN, Yu Y, Xu JT, Li YM, Cai XX, Zhou ZY, Chen XH, Pei YY, Hu L, Su JJ, Cui SD, Wang F, Xie YY, Ding SY, Luo MF, Chou CH, Chang NW, Chen KW, Cheng YH, Wan XH, Hsu WL, Lee TY, Wei FX, Huang HD (2020) miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. Nucleic Acids Res 48:D148–D154

Ison J, Ienasescu H, Chmura P, Rydza E, Ménager H, Kalaš M, Schwämmle V, Grüning B, Beard N, Lopez R (2019) The bio. tools registry of software tools and data resources for the life sciences. Genome Biol 20:1–4

Jiang Q, Wang Y, Hao Y, Juan L, Teng M, Zhang X, Li M, Wang G, Liu Y (2009) miR2Disease: a manually curated database for microRNA deregulation in human disease. Nucleic Acids Res 37: D98–D104

Kalvari I, Argasinska J, Quinones-Olvera N, Nawrocki EP, Rivas E, Eddy SR, Bateman A, Finn RD, Petrov AI (2018) Rfam 13.0: shifting to a genome-centric resource for non-coding RNA families. Nucleic Acids Res 46:D335–D342

Kamanu TK, Radovanovic A, Archer JA, Bajic VB (2013) Exploration of miRNA families for hypotheses generation. Sci Rep 3:1–8

Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I, Papadimitriou D, Kavakiotis I, Maniou S, Skoufos G, Vergoulis T, Dalamagas T, Hatzigeorgiou AG (2018) DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. Nucleic Acids Res 46:D239–D245

Kehl T, Kern F, Backes C, Fehlmann T, Stockel D, Meese E, Lenhof HP, Keller A (2020) miRPathDB 2.0: a novel release of the miRNA pathway dictionary database. Nucleic Acids Res 48:D142–D147

Kozomara A, Griffiths-Jones S (2011) miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 39:D152–D157

Kozomara A, Birgaoanu M, Griffiths-Jones S (2019) miRBase: from microRNA sequences to function. Nucleic Acids Res 47:D155–D162

Kuenne C, Preussner J, Herzog M, Braun T, Looso M (2014) MIRPIPE: quantification of microRNAs in niche model organisms. Bioinformatics 30:3412–3413

Kyrollos DG, Reid B, Dick K, Green JR (2020) RPmirDIP: reciprocal perspective improves targeting prediction. Sci Rep 10:11770

Li J-H, Liu S, Zhou H, Qu L-H, Yang J-H (2014) starBase v2. 0: decoding miRNA-ceRNA, miRNA-ncRNA and protein–RNA interaction networks from large-scale CLIP-Seq data. Nucleic Acids Res 42:D92–D97

Liu CJ, Fu X, Xia M, Zhang Q, Gu Z, Guo AY (2021) miRNASNP-v3: a comprehensive database for SNPs and disease-related variations in miRNAs and miRNA targets. Nucleic Acids Res 49: D1276–D1281

Loher P, Rigoutsos I (2012) Interactive exploration of RNA22 microRNA target predictions. Bioinformatics 28:3322–3323

Lukasik A, Zielenkiewicz P (2019) An overview of miRNA and miRNA target analysis tools. Methods Mol Biol 1932:65–87

Lukasik A, Wojcikowski M, Zielenkiewicz P (2016) Tools4miRs - one place to gather all the tools for miRNA analysis. Bioinformatics 32:2722–2724

Mahlab-Aviv S, Linial N, Linial M (2019) A cell-based probabilistic approach unveils the concerted action of miRNAs. PLoS Comput Biol 15:e1007204

Maselli V, Di Bernardo D, Banfi S (2008) CoGemiR: a comparative genomics microRNA database. BMC Genomics 9:457

Mendes ND, Freitas AT, Sagot MF (2009) Current tools for the identification of miRNA genes and their targets. Nucleic Acids Res 37:2419–2433

Militello G, Weirick T, John D, Döring C, Dimmeler S, Uchida S (2017) Screening and validation of lncRNAs and circRNAs as miRNA sponges. Brief Bioinform 18:780–788

Monga I, Kumar M (2019) Computational resources for prediction and analysis of functional miRNA and their targetome. Methods Mol Biol 1912:215–250

Morel L, Regan M, Higashimori H, Ng SK, Esau C, Vidensky S, Rothstein J, Yang Y (2013) Neuronal exosomal miRNA-dependent translational regulation of astroglial glutamate transporter GLT1. J Biol Chem 288:7105–7116

Müller S, Rycak L, Afonso-Grunz F, Winter P, Zawada AM, Damrath E, Scheider J, Schmäh J, Koch I, Kahl G (2014) APADB: a database for alternative polyadenylation and microRNA regulation events. Database

Naamati G, Friedman Y, Balaga O, Linial M (2012) Susceptibility of the human pathways graphs to fragmentation by small sets of microRNAs. Bioinformatics 28:983–990

Panwar B, Omenn GS, Guan Y (2017) miRmine: a database of human miRNA expression profiles. Bioinformatics 33:1554–1560

Paraskevopoulou MD, Georgakilas G, Kostoulas N, Vlachos IS, Vergoulis T, Reczko M, Filippidis C, Dalamagas T, Hatzigeorgiou AG (2013) DIANA-microT web server v5.0: service integration into miRNA functional analysis workflows. Nucleic Acids Res 41:W169–W173

Paraskevopoulou MD, Vlachos IS, Karagkouni D, Georgakilas G, Kanellos I, Vergoulis T, Zagganas K, Tsanakas P, Floros E, Dalamagas T (2016) DIANA-LncBase v2: indexing microRNA targets on non-coding transcripts. Nucleic Acids Res 44:D231–D238

Perdikopanis N, Georgakilas GK, Grigoriadis D, Pierros V, Kavakiotis I, Alexiou P, Hatzigeorgiou A (2021) DIANA-miRGen v4: indexing promoters and regulators for more than 1500 microRNAs. Nucleic Acids Res 49:D151–D159

Peterson SM, Thompson JA, Ufkin ML, Sathyanarayana P, Liaw L, Congdon CB (2014) Common features of microRNA target prediction tools. Front Genet 5

Preusse M, Theis FJ, Mueller NS (2016) miTALOS v2: analyzing tissue specific microRNA function. PLoS One 11:e0151771

Re A, Caselle M, Bussolino F (2017) MicroRNA-mediated regulatory circuits: outlook and perspectives. Phys Biol 14:045001

Riffo-Campos AL, Riquelme I, Brebi-Mieville P (2016) Tools for sequence-based miRNA target prediction: what to choose? Int J Mol Sci 17

Ru Y, Kechris KJ, Tabakoff B, Hoffman P, Radcliffe RA, Bowler R, Mahaffey S, Rossi S, Calin GA, Bemis L, Theodorescu D (2014) The multiMiR R package and database: integration of microRNA-target interactions along with their disease and drug associations. Nucleic Acids Res 42:e133

Rueda A, Barturen G, Lebron R, Gomez-Martin C, Alganza A, Oliver JL, Hackenberg M (2015) smRNAtoolbox: an integrated collection of small RNA research tools. Nucleic Acids Res 43: W467–W473

Russo F, Di Bella S, Vannini F, Berti G, Scoyni F, Cook HV, Santos A, Nigita G, Bonnici V, Lagana A, Geraci F, Pulvirenti A, Giugno R, De Masi F, Belling K, Jensen LJ, Brunak S, Pellegrini M, Ferro A (2018) miRandola 2017: a curated knowledge base of non-invasive biomarkers. Nucleic Acids Res 46:D354–D359

Sarver AL, Sarver AE, Yuan C, Subramanian S (2018) OMCD: Oncomir cancer database. BMC Cancer 18:1–6

Sato F, Tsuchiya S, Meltzer SJ, Shimizu K (2011) MicroRNAs and epigenetics. FEBS J 278:1598–1609

Schmitz U, Wolkenhauer O (2013) Web resources for microRNA research. Adv Exp Med Biol 774: 225–250

Shaker F, Nikravesh A, Arezumand R, Aghaee-Bakhtiari SH (2020) Web-based tools for miRNA studies analysis. Comput Biol Med 127:104060

Shirdel EA, Xie W, Mak TW, Jurisica I (2011) NAViGaTing the micronome—using multiple microRNA prediction databases to identify signalling pathway-associated microRNAs. PLoS One 6:e17429

Shuang C, Maozu G, Chunyu W, Xiaoyan L, Yang L, Xuejian W (2016) MiRTDL: a deep learning approach for miRNA target prediction. IEEE/ACM Trans Comput Biol Bioinform 13:1161–1169

Shukla V, Varghese VK, Kabekkodu SP, Mallya S, Satyamoorthy K (2017) A compilation of web-based research tools for miRNA analysis. Brief Funct Genomics 16:249–273

Solomon J, Kern F, Fehlmann T, Meese E, Keller A (2020) HumiR: web services, tools and databases for exploring human microRNA data. Biomolecules 10

Sticht C, De La Torre C, Parveen A, Gretz N (2018) miRWalk: an online resource for prediction of microRNA binding sites. PLoS One 13:e0206239

The RNAcentral Consortium, Petrov AI, Kay SJE, Kalvari I, Howe KL, Gray KA, Bruford EA, Kersey PJ, Cochrane G, Finn RD, Bateman A, Kozomara A, Griffiths-Jones S, Frankish A, Zwieb CW, Lau BY, Williams KP, Chan PP, Lowe TM, Cannone JJ, Gutell R, Machnicka MA, Bujnicki JM, Yoshihama M, Kenmochi N, Chai B, Cole JR, Szymanski M, Karlowski WM, Wood V, Huala E, Berardini TZ, Zhao Y, Chen R, Zhu W, Paraskevopoulou MD, Vlachos IS, Hatzigeorgiou AG, Ma L, Zhang Z, Puetz J, Stadler PF, McDonald D, Basu S, Fey P, Engel SR, Cherry JM, Volders P-J, Mestdagh P, Wower J, Clark MB, Quek XC, Dinger ME (2017) RNAcentral: a comprehensive database of non-coding RNA sequences. Nucleic Acids Res 45:D128–D134

Thomson DW, Bracken CP, Goodall GJ (2011) Experimental strategies for microRNA target identification. Nucleic Acids Res 39:6845–6853

Tokar T, Pastrello C, Rossos AE, Abovsky M, Hauschild A-C, Tsay M, Lu R, Jurisica I (2018) mirDIP 4.1—integrative database of human microRNA target predictions. Nucleic Acids Res 46:D360–D370

Tong Y, Ru B, Zhang J (2018) miRNACancerMAP: an integrative web server inferring miRNA regulation network for cancer. Bioinformatics 34:3211–3213

Tong Z, Cui Q, Wang J, Zhou Y (2019) TransmiR v2.0: an updated transcription factor-microRNA regulation database. Nucleic Acids Res 47:D253–D258

Van Peer G, Lefever S, Anckaert J, Beckers A, Rihani A, Van Goethem A, Volders PJ, Zeka F, Ongenaert M, Mestdagh P, Vandesompele J (2014) miRBase tracker: keeping track of microRNA annotation changes. Database (Oxford)

Vejnar CE, Blum M, Zdobnov EM (2013) miRmap web: comprehensive microRNA target prediction online. Nucleic Acids Res 41:W165–W168

Wang P, Zhi H, Zhang Y, Liu Y, Zhang J, Gao Y, Guo M, Ning S, Li X (2015) miRSponge: a manually curated database for experimentally supported miRNA sponges and ceRNAs. Database

Wienholds E, Plasterk RH (2005) MicroRNA function in animal development. FEBS Lett 579:5911–5922

Wong NW, Chen Y, Chen S, Wang X (2018) OncomiR: an online resource for exploring pan-cancer microRNA dysregulation. Bioinformatics 34:713–715

Wu W-S, Tu B-W, Chen T-T, Hou S-W, Tseng JT (2017) CSmiRTar: condition-specific microRNA targets database. PLoS One 12:e0181231

Xiao F, Zuo Z, Cai G, Kang S, Gao X, Li T (2009) miRecords: an integrated resource for microRNA-target interactions. Nucleic Acids Res 37:D105–D110

Xie B, Ding Q, Han H, Wu D (2013) miRCancer: a microRNA–cancer association database constructed by text mining on literature. Bioinformatics 29:638–644

Yang J-H, Li J-H, Jiang S, Zhou H, Qu L-H (2013) ChIPBase: a database for decoding the transcriptional regulation of long non-coding RNA and microRNA genes from ChIP-Seq data. Nucleic Acids Res 41:D177–D187

Yang Z, Wu L, Wang A, Tang W, Zhao Y, Zhao H, Teschendorff AE (2017) dbDEMC 2.0: updated database of differentially expressed miRNAs in human cancers. Nucleic Acids Res 45:D812–D818

You Z-H, Huang Z-A, Zhu Z, Yan G-Y, Li Z-W, Wen Z, Chen X (2017) PBMDA: a novel and effective path-based computational model for miRNA-disease association prediction. PLoS Comput Biol 13:e1005455

Yue D, Liu H, Huang Y (2009) Survey of computational algorithms for MicroRNA target prediction. Curr Genomics 10:478–492

Zou Q, Li J, Hong Q, Lin Z, Wu Y, Shi H, Ju Y (2015) Prediction of microRNA-disease associations based on social network analysis methods. BioMed Res Int

Chapter 6
Bioinformatics Methods for Modeling microRNA Regulatory Networks in Cancer

Tao Pan, Yueying Gao, Gang Xu, and Yongsheng Li

Abstract MicroRNAs (miRNAs) play important roles in the physiology and development of cancers. The increase of multidimensional molecular profiles of tumor patients generated by high-throughput sequencing technologies has enabled computational analysis of miRNA regulatory networks in cancer. In this chapter, we first summarized currently widely used computational methods for identifying miRNA–gene interactions. In addition, crosstalk among miRNAs and competitive endogenous RNAs (ceRNAs) represent novel layers of gene regulation mediated by miRNAs, which also play important roles in cancer. We next reviewed computational methods for modeling miRNA–miRNA crosstalk and ceRNA–ceRNA interactions in cancer. These methods integrate multi-omics data and range from genomics to phenomics. MiRNA–miRNA networks are generally constructed based on genomic sequences, transcriptomes, miRNA–gene regulation, and functional pathways. Moreover, five types of computational methods for identifying ceRNA–ceRNA interactions are summarized in this chapter. Among these methods, two types of global ceRNA regulation and three types of context-specific methods are included. The application of these computational methods focused on miRNA regulation in cancer provides valuable functional insights into the underlying mechanism of cancer, as well as future precision medicine.

Keywords miRNA · Regulatory network · Cancer · ceRNA · Crosstalk

6.1 Introduction

A microRNA (miRNA) is a kind of small endogenous RNA (Lu and Rothenberg 2018), which is about 22 nucleotides in length (Bartel 2009) but does not encode protein. MiRNAs play a crucial role in the process of protein synthesis and function.

T. Pan · Y. Gao · G. Xu · Y. Li (✉)
College of Biomedical Information and Engineering, Hainan Medical University, Haikou, China
e-mail: liyongsheng@hainmc.edu.cn

U. Schmitz et al. (eds.), *Systems Biology of MicroRNAs in Cancer*, Advances in Experimental Medicine and Biology 1385,
https://doi.org/10.1007/978-3-031-08356-3_6

MiRNAs are key regulators of gene selective expression in the process of cell differentiation by targeting mRNAs. In addition, miRNAs also play an important regulatory role in the development and progression of human diseases (Chang and Mendell 2007; Zhang 2008; Xu et al. 2020), such as prostate cancer, breast cancer, and colorectal cancer.

In order to better understand the mechanism of miRNAs and their important biological functions, the key is to recognize the targets of miRNAs. MiRNAs can degrade or inhibit the translation of target mRNAs by binding to RISC (RNA-induced silencing complex) (Bartel 2018; Kobayashi and Tomari 2016) and acting on the 3′-UTR of mRNAs. The "seed sequence" is the 2–8 nucleotide sequence of the 5′-end of miRNAs (Lewis et al. 2003), which can complement and pair with the 3′-UTR of mRNAs. The inhibitory effects of miRNAs on mRNAs depend on the way in which the seed sequence of miRNAs binds with the 3′-UTR of mRNAs. If the seed sequence of a miRNA completely complements and pairs with a 3′-UTR subsequence (a.k.a. miRNA response element) of the target mRNA it is called "seed matching" (Lewis et al. 2003), then miRNA will directly degrade its target mRNA for incomplete pairing. A miRNA can also inhibit the translation of its target mRNA. Large numbers of computational methods have been proposed to identify the targets of miRNAs, such as TargetScan (Agarwal et al. 2015) and DIANA-TarBase (Karagkouni et al. 2018). All these methods provide novel insights into the functions of miRNAs in complex diseases.

Moreover, miRNA–miRNA crosstalk has been found in various kinds of cancers (Shao et al. 2019). However, the identification of miRNA–miRNA crosstalk based on experimental methods is still challenging. Computational methods have dramatically reduced the number of miRNA–miRNA crosstalk candidates. These methods range from genomics to phenomics. On the other side, crosstalk among RNAs yields large regulatory competing endogenous RNAs (ceRNAs) via co-regulated miRNAs. CeRNA regulation represents a novel layer of gene regulation that plays important roles in complex diseases (Xu et al. 2015, 2016). A number of computational methods have been proposed to model ceRNA–ceRNA interaction networks (Li et al. 2018, 2019). Understanding these computational techniques focused on miRNA–miRNA or ceRNA–ceRNA interaction will help understand the miRNA regulatory mechanisms in complex diseases.

6.2 Bioinformatics Methods for miRNA–Gene Regulatory Networks

1. *TargetScan*
   Although the sequence-based algorithms for predicting the targets of miRNAs are different from each other, they basically follow similar principles: (a) The seed region of miRNAs is complementary to the 3′-UTR of mRNAs; (b) The binding of miRNA and mRNA results in thermodynamic stability of the double-stranded

structure; (c) MiRNA targets are conserved in different species; (d) The prediction of miRNA targets is impacted by the secondary structure of miRNA targets and the sequences outside miRNA targets.

TargetScan is a prediction algorithm developed in 2003 to predict and recognize the miRNA targets (Agarwal et al. 2015). The algorithm depending on sequencing data predicts the target genes of miRNAs, and relies on a 3′-UTR configuration file. Two concepts of "miRNA seed" and "seed matching" were proposed for the first time. TargetScan predicts the miRNA targets by searching for conserved 8-mer, 7-mer, and 6-mer sites ("mer" refers to the length of nucleic acid sequence) that are matched to each miRNA seed region (Fig. 6.1). For this tool, types of binding site and another 14 features are considered to identify the most effective miRNA targets.

In 2018, TargetScan released version 7.2, which covers miRNA target information in more than ten species including humans and mice (Fig. 6.2). It includes 19,432 human genes and corresponding 28,352 transcripts, and 2,606 human miRNAs. All mature miRNAs are obtained from the miRbase database. TargetScan classifies miRNAs based on their conservation, and users can select miRNAs from different species according to different classifications. The output shows all mRNAs containing the binding sites of interested miRNA and provides the gene name of miRNA targets and the corresponding transcripts, a number of 3P-seq tags supporting UTR+5, binding site on UTR, the number of bases paired with the seed sequence in the conservative and nonconservative binding sites, and the score of the binding sites (Fig. 6.3). The current TargetScan examined 3′-UTR isoform quantifications previously generated using poly(A)-position profiling by sequencing (3P-seq). The model used the dominant mRNA from the subset of genes for which $\geq$90% of the 3P-seq tags corresponded to a single 3′-UTR isoform. The number of 3P-seq tags supporting UTR indicates this information. TargetScan returns a score for each binding site according to the context++ model, which is integrated by the 14 features of the specific binding site. The smaller the score, the greater the probability of being a miRNA target. TargetScan also provides another parameter, aggregate Pct, which is used to evaluate the conservation of miRNA binding sites. The higher the aggregate Pct score, the higher the conservation of the miRNA binding sites.

2. *DIANA-TarBase*

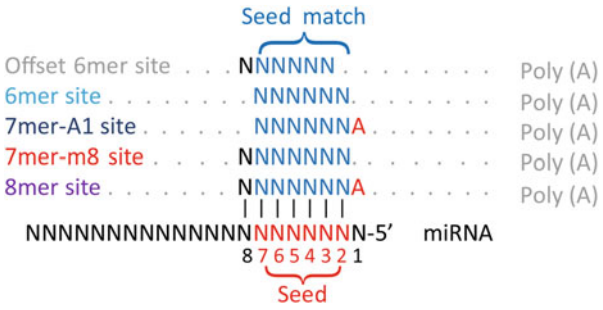


Fig. 6.1 The sites of interactions between miRNAs and mRNAs

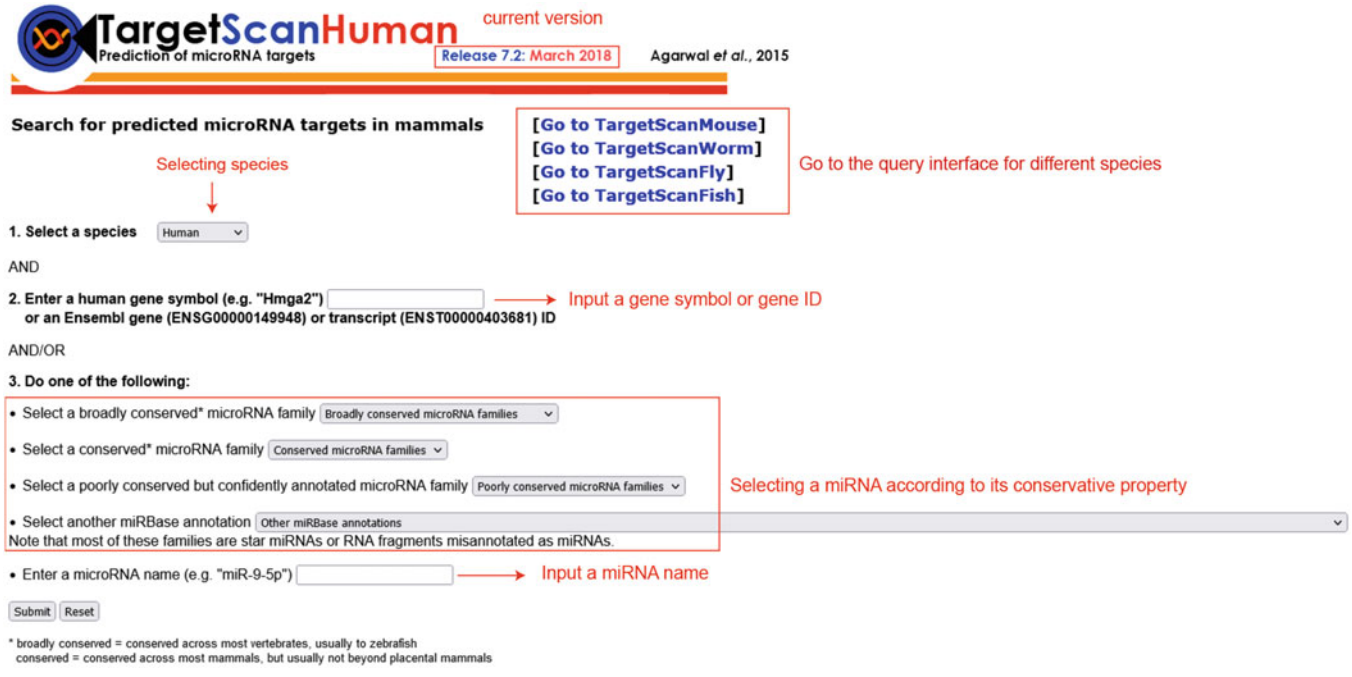
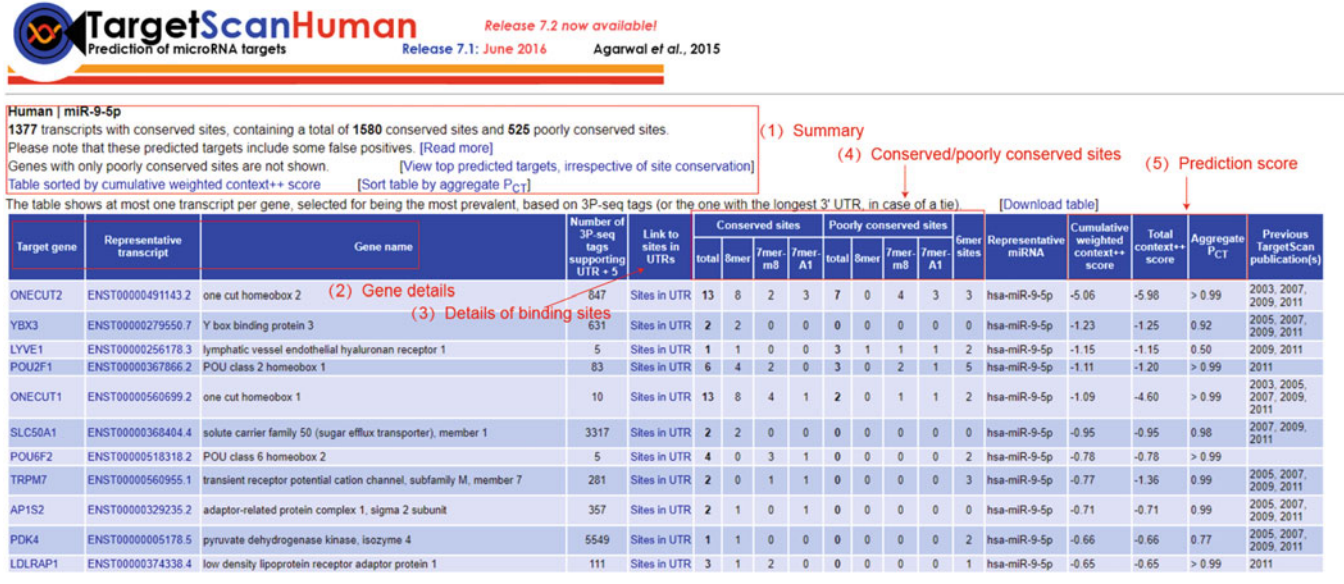


Fig. 6.2 Home page of TargetScan



Target gene	Representative transcript	Gene name	Number of 3P-seq tags supporting UTR + 5	Link to sites in UTRs	Conserved sites				Poorly conserved sites				6mer sites	Representative miRNA	Cumulative weighted context++ score	Total context++ score	Aggregate P_{CT}	Previous TargetScan publication(s)
					total	8mer	7mer-m8	7mer-A1	total	8mer	7mer-m8	7mer-A1						
ONECUT2	ENST00000491143.2	one cut homeobox 2	847	Sites in UTR	13	8	2	3	7	0	4	3	3	hsa-miR-9-5p	-5.06	-5.98	> 0.99	2003, 2007, 2009, 2011
YBX3	ENST00000279550.7	Y box binding protein 3	631	Sites in UTR	2	2	0	0	0	0	0	0	0	hsa-miR-9-5p	-1.23	-1.25	0.92	2005, 2007, 2009, 2011
LYVE1	ENST00000256178.3	lymphatic vessel endothelial hyaluronan receptor 1	5	Sites in UTR	1	1	0	0	3	1	1	1	2	hsa-miR-9-5p	-1.15	-1.15	0.50	2009, 2011
POU2F1	ENST00000367866.2	POU class 2 homeobox 1	83	Sites in UTR	6	4	2	0	3	0	2	1	5	hsa-miR-9-5p	-1.11	-1.20	> 0.99	2011
ONECUT1	ENST00000560699.2	one cut homeobox 1	10	Sites in UTR	13	8	4	1	2	0	1	1	2	hsa-miR-9-5p	-1.09	-4.60	> 0.99	2003, 2005, 2007, 2009, 2011
SLC50A1	ENST00000368404.4	solute carrier family 50 (sugar efflux transporter), member 1	3317	Sites in UTR	2	2	0	0	0	0	0	0	0	hsa-miR-9-5p	-0.95	-0.95	0.98	2007, 2009, 2011
POU6F2	ENST00000518318.2	POU class 6 homeobox 2	5	Sites in UTR	4	0	3	1	0	0	0	0	2	hsa-miR-9-5p	-0.78	-0.78	> 0.99	
TRPM7	ENST00000560955.1	transient receptor potential cation channel, subfamily M, member 7	281	Sites in UTR	2	0	1	1	0	0	0	0	3	hsa-miR-9-5p	-0.77	-1.36	0.99	2005, 2007, 2009, 2011
AP1S2	ENST00000329235.2	adaptor-related protein complex 1, sigma 2 subunit	357	Sites in UTR	2	1	0	1	0	0	0	0	0	hsa-miR-9-5p	-0.71	-0.71	0.99	2005, 2007, 2009, 2011
PDK4	ENST00000005178.5	pyruvate dehydrogenase kinase, isozyme 4	5549	Sites in UTR	1	1	0	0	0	0	0	0	2	hsa-miR-9-5p	-0.66	-0.66	0.77	2005, 2007, 2009, 2011
LDLRAP1	ENST00000374338.4	low density lipoprotein receptor adaptor protein 1	111	Sites in UTR	3	1	2	0	0	0	0	0	1	hsa-miR-9-5p	-0.65	-0.65	> 0.99	2011

Fig. 6.3 Results page of TargetScan

DIANA-TarBase v8 was released by the DIANA laboratory in 2018 (Karagkouni et al. 2018). TarBase is one of the databases for querying the targets of miRNAs confirmed by experiments (Fig. 6.4). The database contains the interactions between miRNAs and their targets in 18 species. There are a total of more than 1.08 million entries, which were generated by 34 methods, spanning 85 tissue types, and reported by 1208 articles. The web address is: http://carolina.imis.athena-innovation.gr/diana_tools/web/index.php?r=tarbasev8/index.

The functionality of TarBase is very powerful and practical. In the home page of the database, there is a prominet statistics panel for users. By clicking this icon, we can see statistical results of all the information in this database, including species, cell types, and source of tissues. The database provides several options for users, such as species, method type, method, regulation type, and validation type. The database also provides information from different versions of TarBase for users to query, including TarBase 6.0 (Vergoulis et al. 2012), TarBase 7.0 (Vlachos et al. 2015a), and TarBase 8.0. For the miRNA queried in this database, the name that conforms to the miRNA naming format in the miRBase database should be provided, such as hsa-let-7a-3p. However, for genes, we can input the gene symbol or Ensemble ID.

This database can return a comprehensive information for users' querying (Fig. 6.5). For target genes, we can observe the chromosome location, transcript ID, gene type (coding gene or noncoding gene), Ensemble gene ID (by clicking, it can also link to external database to obtain the details), gene names, and functional parts. For miRNAs, the query results include the name of the miRNA, the sequence, miRBase ID (link to an external database to obtain the detailed information on miRNA), the information of the miRNA in other tools of DIANA, and diseases associated with the miRNA. Similarly, in the result interface, the tool also provides filtering options. However, the data in the database can only be browsed, but cannot be downloaded. If you want to download, you need to send an application to the development team.

3. *PicTar*

PicTar (probabilistic identification of combinations of target sites) predicts the miRNA targets by constructing a probabilistic model (Krek et al. 2005). PicTar algorithm can not only effectively predict the targets of a single miRNA, but also predict the common targets of multiple miRNAs. At the same time, it can reduce the false positive rate caused by the small number of miRNA binding sites on mRNA 3′-UTR.

The input file of PicTar is a miRNA set and a file of homologous 3′-UTR sequences. PicTar calculates the maximum probability that all binding sites on the sequence bind to each miRNA in the set, and filters the false positives through cross-species comparison. Then PicTar integrates the scores of candidate sequences of each species, and finally gets the PicTar score of genes (Fig. 6.6). PicTar's prediction performance has been well verified in the prediction of the targets of multiple miRNAs across species, and the false positive has been constantly low. It is the first algorithm to predict common targets of miRNA

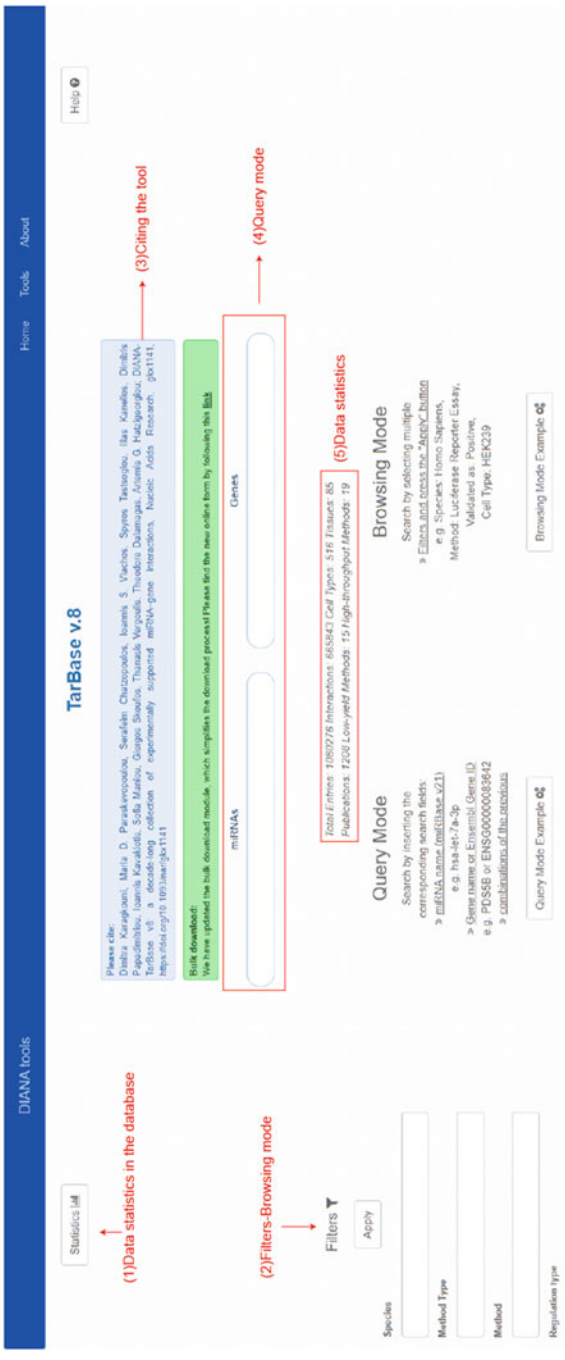


Fig. 6.4 Home page of TarBase

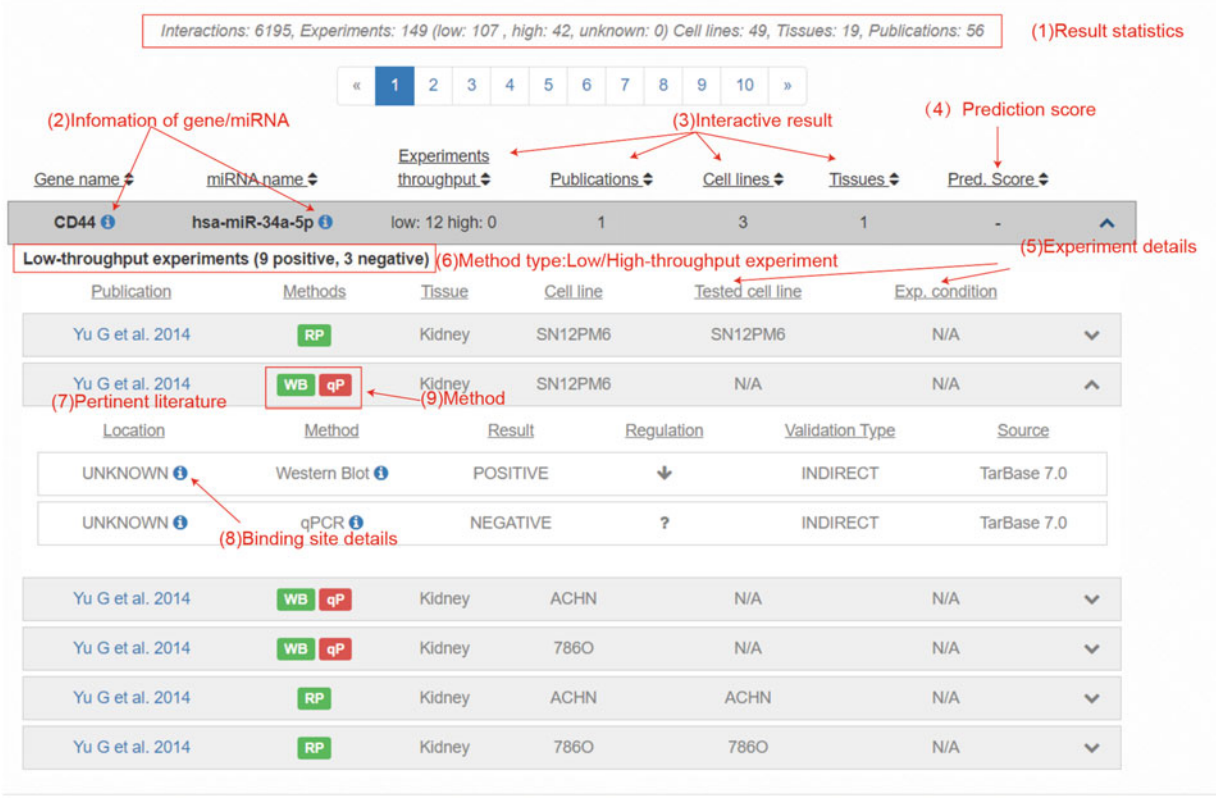


Fig. 6.5 Result page for hsa-miR-34a-5p in TarBase

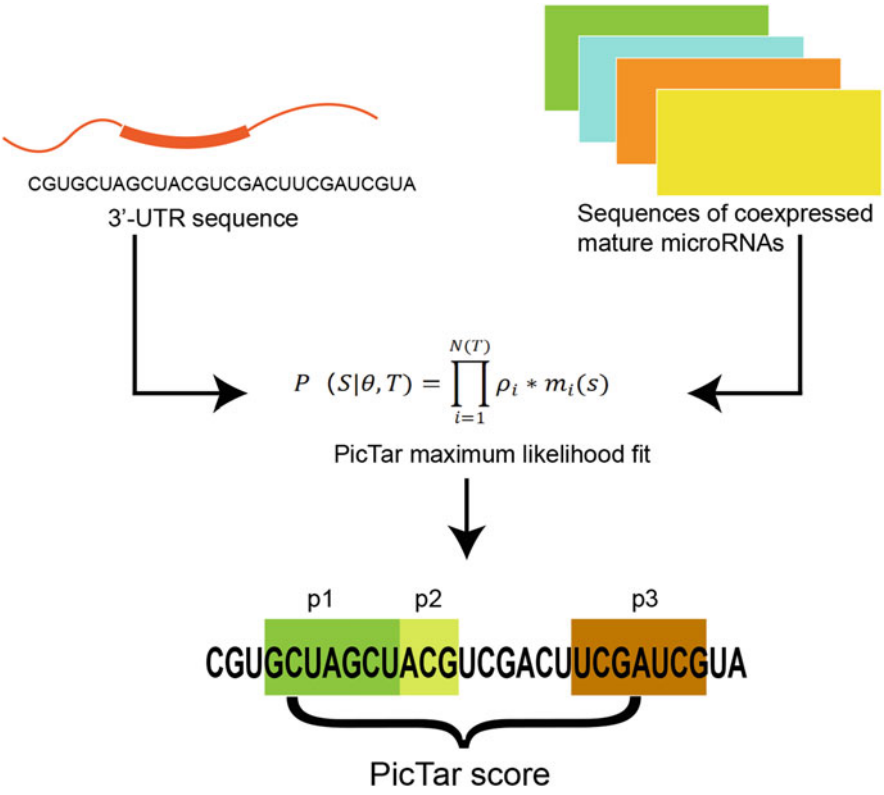


Fig. 6.6 Flowchart for calculating the PicTar score

collections. PicTar analysis results are available from https://pictar.mdc-berlin.de/.

4. *HOCTAR*

HOCTAR is another algorithm that predicts miRNA targets based on miRNA and mRNA expression profiles (Fig. 6.7). The algorithm predicts miRNA targets by expression of human miRNA host genes, and has been applied to 178 human miRNA host genes (Gennarino et al. 2009). The algorithm obtains the information of miRNA in human gene and its corresponding host gene from miRBase version 10.1. The g:Sorter tool (http://biit.cs.ut.ee/gprofiler/gost) is used to infer the expression correlation between miRNA and its corresponding host genes. The algorithm is also compared with other miRNA target gene prediction tools, such as TargetScan, Miranda, PicTar, and other tools, and it proves that the accuracy of HOCTAR in predicting miRNA targets is reliable. The results of prediction analysis are available at: http://hoctar.tigem.it.

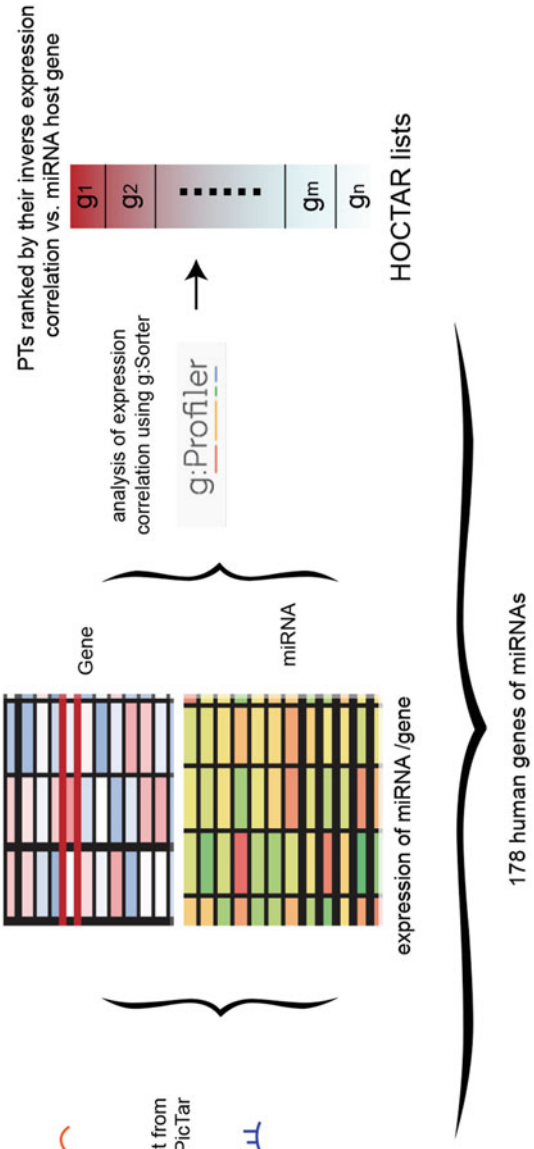


Fig. 6.7 MiRNA targets prediction workflow of HOCTAR

6.3 Bioinformatics Methods for Analyzing miRNA–miRNA Regulatory Networks

6.3.1 Sequence Similarity: Complementary Base Pairing

Different miRNA sequences can bind to each other in a complementary way, thereby affecting the transcription of each other (Fig. 6.8a). Guo et al. (2012) analyzed the expression patterns of miRNAs in different species, and found miRNA–miRNA complementary matching in different species (such as human, rat, and mouse). Some of the miRNA pairs are relatively well conserved among species.

In the past few decades, a number of bioinformatics resources, methods, and tools had been developed to predict and store the miRNA–target interactions (Table 6.1). These tools predict miRNA–target interactions relatively reliable by evaluating the accessibility, conservation, stability, correlation, and complementarity of the sites and the scores obtained by the developed algorithms. In addition, miRBase database (Kozomara et al. 2019) stores pre-miRNA and mature miRNAs sequences of many animal species. We can obtain endogenous sense and antisense miRNAs according to the specific location of these miRNAs in the genome, and then obtain a pair of miRNA–miRNA complementary matching. These databases contribute to investigating the interactions among sequence-based miRNAs. Therefore, by analyzing the similarity of seed sequences among miRNAs, we can get a global view of miRNA crosstalk.

6.3.2 Higher-Order Chromatin Conformation

With the emergence and development of techniques to capture chromosome conformation, it has been proved that chromatin interactions play a role in transcriptional regulation and coordination, and affect genomic function (Jia et al. 2017). Furthermore, we have studied the influence of higher chromatin structures on the regulation of miRNA transcription (Fig. 6.8b). Chen et al. (2014) investigated how higher-order chromatin structures regulate miRNA transcription, and found that when miRNA and protein-coding genes were in the same position, there would be coordinate expression. They also linked highly reliable miRNA–target pairs with chromatin interaction networks, further suggesting the existence of spatial miRNA–miRNA chromatin interaction networks, and indicated that in general, spatially coordinated miRNA pairs were from the same family and related to the same disease category.

As we study chromatin interactions in more depth, the data about chromatin interactions has increased rapidly (Table 6.1). Teng et al. (2015) constructed the 4DGenome database for storing chromatin interactions through a comprehensive literature survey. The database contains the datasets obtained by low and high flux measurement methods including 3C, 4C, 5C, Hi-C, and so on. Xie et al. (2016) also

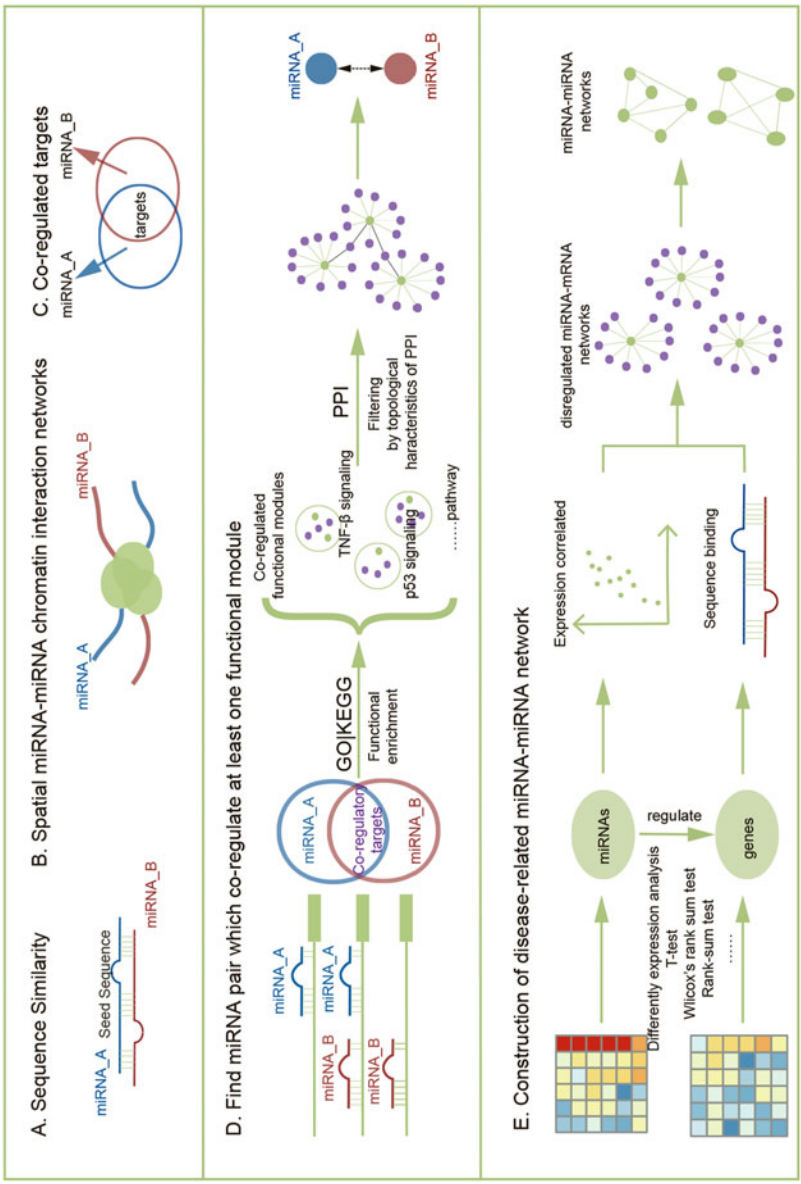


Fig. 6.8 Methods based on various omics to construct the miRNA–miRNA functional co-regulation networks. (**a**) Methods based on sequence similarity; (**b**) Methods based on chromatin interaction networks; (**c**) Methods based on co-regulated targets; (**d**) Methods based on functional modules and protein–protein interactions; (**e**) Methods based on disease-related dysregulated miRNAs and genes

Table 6.1 Commonly used methods to construct the miRNA–miRNA network

Methods	Tool	PubMed ID	Website
Sequence similarity			
Guo et al.	×	23031806	–
TargetScan	√	26267216	http://www.targetscan.org/vert_72/
miRnada	√	20799968	http://cbio.mskcc.org/microrna_data/miRanda-aug2010.tar.gz
Starbase	√	24297251	http://starbase.sysu.edu.cn/
miRBase	√	30423142	http://mirbase.org/
Higher-order chromatin conformation			
Jia et al.	×	29149895	–
Chen et al.	×	24357409	–
4DGenome	√	25788621	http://4dgenome.int-med.uiowa.edu/
Xi et al.	√	26868054	http://songyanglab.sysu.edu.cn/ccsi
Co-regulated genes			
Fan et al.	×	25471818	–
miRTarBase	√	31647101	http://miRTarBase.cuhk.edu.cn/
miRDB	√	31504780	http://mirdb.org/
miRBase	√	30423142	http://mirbase.org/
miRWalk	√	26226356	http://mirwalk.umm.uni-heidelberg.de/
miRmapper	√	30223528	R package miRmapper
DIANA miRPath v.2.0	√	22649059	http://www.microrna.gr/miRPathv2
DIANA-miRPath v3.0	√	25977294	http://www.microrna.gr/miRPathv3
Xu et al.	×	27551063	–
Shalgi et al.	×	17630826	–
Segura et al.	×	25838464	–
GeneSet2miRNA	√	19420064	http://mips.helmholtz-muenchen.de/proj/gene2mir/
miRror	√	20529892	http://www.proto.cs.huji.ac.il/mirror/
Functional similarity			
C2Analyzer	√	24862384	http://www.bioinformatics.org/c2analyzer/
Aftabuddin et al.	×	26066638	–
Hsu et al.	×	18491312	–
Liang et al.	×	17652130	–
Sun et al.	×	23874989	–
Xu et al.	×	20929877	–
Disease phenotype			
HMDD	√	30364956	http://www.cuilab.cn/hmdd
Chaulk et al.	×	26563430	–
Hua et al.	×	25641175	–
Hua et al.	×	23619378	–
Ting et al.	×	29800060	–
miR2Disease	√	18927107	http://www.mir2disease.org/
CancerNet	√	26690544	http://bis.zju.edu.cn/CancerNet

acquired 91 sets of chromatin interaction data from published literature. Through the integration and annotation of datasets, a database of chromatin–chromatin spatial interaction (CCSI) was constructed. These chromatin datasets provide us with resources to study the interactions among ncRNAs (non-coding RNA, including miRNAs) at the chromatin level.

6.3.3 Co-regulated Genes

To detect the interactions among miRNAs, we can proceed from miRNA targets. If miRNA pairs can co-regulate at least one target together, they may cooperate with each other (Fig. 6.8c). miRWalk2.0 (Dweep and Gretz 2015) is a comprehensive database that integrates information of miRNA–target interactions derived from miRTarBase, miRDB, TargetScan, and other resources. The database also provides an additional framework to obtain miRNA pairs that can significantly co-regulate genes. In the investigation of Silveira et al. (da Silveira et al. 2018), the authors developed miRmapper, which could identify the main miRNAs in the miRNA–mRNA regulatory network and calculated the similarity among miRNAs using Jaccard distance based on the co-regulatory genes, further revealed the cooperative interactions of miRNAs in the regulation of common targets. DIANA-miRPath v.2.0 (Vlachos et al. 2012) proposed an analysis scheme for the intersection or union of genes (or pathways) to study miRNA pairs, while DIANA-miRPath v3.0 (Vlachos et al. 2015b) was updated by adding functional annotations to identify miRNAs in specific GO terms or pathways.

In current studies related to identifying cooperative miRNA target regulation, the common statistical test is the cumulative hypergeometric statistic. Some of the above databases have analyzed the significance of miRNA targets by using the hypergeometric test. However, there may be potential bias when using the hypergeometric distribution to predict cooperative miRNA regulation because of variation in the average length of the 3′-UTR. Targets with a longer 3′-UTR may contain more miRNA binding sites, which may lead to an overrated co-occurrence rate (very low *P*-value) (Xu et al. 2017). For the scheme of inferring interactions in the miRNA–gene regulatory network, Shalgi et al. (2007) designed a random-based test. With regards to random miRNA–gene regulation, the Meet/Min score of each miRNA–miRNA pair was calculated. The co-regulation P-value was defined according to the Meet/Min score of the pair and used to identify the co-regulated miRNA pairs. Segura et al. (2015) also extended the clustering coefficient, which was used to describe the degree of clustering between the vertices of a graph, into the network by using the Meet/Min coefficient, greatly increasing the accuracy of the predicted interaction.

In addition to co-regulated genes, there are also many methods to identify cooperative miRNA pairs via the list of co-regulated genes. The GeneSet2miRNA database (Antonov et al. 2009) is a network tool to identify the miRNA activity by gene lists. The input of this tool is a gene list. It uses the hypergeometric test and

some other tests to identify miRNA models (single, pair, triplet, or quadruplet miRNAs) significantly related to the gene set, and finally outputs the list of miRNA regulatory models. Mirror (Friedman et al. 2010) also used gene sets to explain the possibility of the observed data by evaluating the combined regulation of miRNA targets under various conditions, and finally found the best miRNA set to explain its regulation. These tools based on gene sets have identified the crosstalk among miRNAs, but the gene set used also limited these tools.

6.3.4 Functional Similarity

Recent studies have demonstrated the prediction of miRNA pairs based on their functions. It is assumed that genes regulated by miRNA combinations may have similar functions and play similar roles. According to the above hypothesis, many methods have been proposed to identify miRNA pairs with similar functions by evaluating their functional similarity. C2Analyzer (Co-target-Co-function Analyzer) (Aftabuddin et al. 2014) is a multifunctional and user-friendly web tool based on Perl. In order to determine whether miRNA pairs are functionally enriched, this tool uses hypergeometric analysis. It can also output miRNA result files, which is convenient for us to visualize miRNA–miRNA pairs by using software such as Cytoscape. Aftabuddin et al. (Mal et al. 2015) have identified the functional cooperative module of miRNA pairs and constructed a miRNA–miRNA functional synergy network by combining the co-regulatory target network with Gene Ontology (GO) annotation. Functional similarity usually depends on whether co-regulatory genes share the same GO annotations or pathways, which is usually identified by the hypergeometric test. Though proteins regulated by miRNAs might not form regulatory modules, miRNA–target genes can form significant regulatory modules by binding with their interacting proteins (Hsu et al. 2008).

Previous studies have shown that interacting proteins tend to be regulated by similar types of miRNAs (Liang and Li 2007). Sun et al. (2013) developed a new computing framework and method, miRFunSim, based on graph properties. This tool calculated the functional similarity scores of miRNA pairs according to the protein connectivity in the protein–protein interaction network and the targeting orientation of the miRNAs. It also revealed that the functional similarity scores of miRNAs in the same family or cluster were significantly higher than other miRNAs, which was confirmed by results of previous studies. Xu et al. (2011) developed a computational method to construct a miRNA–miRNA functional synergetic network. By integrating three characteristics, (1) the common targets of miRNA pairs, (2) the enrichment of the same GO terms, and (3) the topological characteristics of the protein–protein interaction network, they obtained the functional module of common regulation, and then identified the synergetic miRNA pairs based on this module to construct a miRNA–miRNA functional synergetic network (Fig. 6.8d).

6.3.5 Disease Phenotype

The above part is based on the global identification of miRNA pairs, using sequence, co-regulatory genes, functions, and other information. However, if we want to construct a miRNA–miRNA network for a certain disease, we need to integrate more information and consider more aspects, such as miRNAs associated with human diseases, expression profile of miRNAs, and their regulated genes in specific disease, genes regulated by miRNAs in specific disease and so on.

HMDD (Huang et al. 2019) is a comprehensive database of human disease-related miRNAs that integrates a large number of miRNA–disease associated entries from the published literature and is constantly being maintained and updated. This database is conducive to investigate miRNAs related to diseases. Chaulk et al. (2016) investigated whether there were co-expression patterns among miRNAs using miRNA expression data, and found complex miRNA co-expression patterns and determined that co-expressed miRNAs had similar biological activities. We can also integrate the differentially expressed miRNAs with other information to construct a miRNA–miRNA network (Fig. 6.8e). Hua et al. (2014) built a synergistic miRNA–miRNA network related to coronary artery disease by combining the differentially expressed miRNAs and genome-wide single nucleotide polymorphism (SNP) genotype. After determining the regulating relationship between miRNAs and mRNAs in a specific environment, above methods are used to construct miRNA–miRNA interaction networks.

Hua et al. (2013) used the expression data of miRNAs and mRNAs in breast cancer to construct the maladjusted miRNA–mRNA network. Then based on this network, the miRNA–miRNA network was extracted according to the correlation coefficient. They also applied other network analysis methods to the miRNA expression data in order to verify the identified miRNA clusters. In addition to investigations on specific cancers, there are studies that analyze multiple types of cancer. Ting et al. (Shao et al. 2019) studied the collaborative regulation of miRNA pairs in 18 cancer types, and found that cancer types with similar tissue origin had high similarity in the expression of the collaborative network and collaborative miRNA pairs. They further identified miRNA hubs and combined miRNA collaboration modules with clinical information for survival analysis, and found that miRNA collaboration modules were related to the survival rate of patients with several types of cancer. Meng et al. (2015) developed the CancerNet database and respectively constructed cancer-specific miRNA–miRNA synergistic networks based on the functions of miRNA targets and their topological characteristics in a cancer protein interactions network for 33 human cancers. CancerNet can serve as a comprehensive platform for evaluating the interactions between proteins and miRNAs in human cancers.

6.4 Bioinformatics Methods for Identifying miRNA-Mediated ceRNA Networks

An increasing body of evidence suggests that >60% of protein-coding genes (PCGs) are regulated by miRNAs (Bajan and Hutvagner 2014; Friedman et al. 2009). Moreover, many PCGs can also be regulated by several miRNAs. Pandolfi et al. found that some protein–coding genes and their pseudogenes had the same miRNA-binding sites in their 3′-UTRs, and that their respective expression levels were regulated by competing for miRNA binding (Poliseno et al. 2010). Based on this mechanism, they proposed the competing endogenous RNA (ceRNA) hypothesis (Salmena et al. 2011). According to the hypothesis, ceRNAs could act as molecular sponges for a miRNA through their miRNA response elements (MREs) to regulate other target genes of the respective miRNAs. The conventional regulatory pattern has been altered from linear regulation of miRNA–RNA to network regulation of RNA–miRNA–RNA. Understanding this novel type of RNA regulatory pattern will provide insights into regulatory networks and may help develop strategies in human cancer and other complex diseases (Li et al. 2019). The ceRNA hypothesis has aroused a wide concern and various types of RNAs, including pseudogenes, lncRNAs, and circRNAs, as well as mRNAs were demonstrated to be ceRNA molecules (Fig. 6.9). On the basis of this hypothesis, MREs can be viewed as the letters of an "RNA language", and transcripts can actively communicate with others to regulate their respective expression levels (Fig. 6.9).

After obtaining the miRNA–target regulation, there are two commonly used principles to identify miRNA-mediated ceRNA regulatory networks (Ala et al. 2013; Chiu et al. 2015a). The central hypothesis of most computational methods is that ceRNA crosstalk increases with high miRNA regulatory similarity between mRNAs and their strong co-expression in a specific context (Fig. 6.10) (Li et al. 2019).

Firstly, miRNA–target interactions were discovered by integration of computational methods and AGO-CLIP Seq datasets. Secondly, miRNA-mediated ceRNA pairs are identified by ratio or hypergeometric test. Finally, RNA–RNA expression similarity was evaluated based on expression profile datasets.

Here, we reviewed five types of methods for identifying ceRNA regulation or miRNA sponge interactions (Table 6.2), including two types of global ceRNA regulation prediction methods (ratio-based, termed ratio, and hypergeometric test-based, termed HyperT) and three types of context-specific prediction methods (Hypergeometric test combined with co-expression, termed HyperC, sensitivity correlation (SC)-based method and conditional mutual information (CMI)-based methods). Collectively, these strategies offer a major advantage in identifying the interactions of functional miRNA–targets.

For a ceRNA pair of RNA-X and RNA-Y, the following four types of methods are used to estimate whether RNA-Y is a modulator of RNA-X (Fig. 6.11).

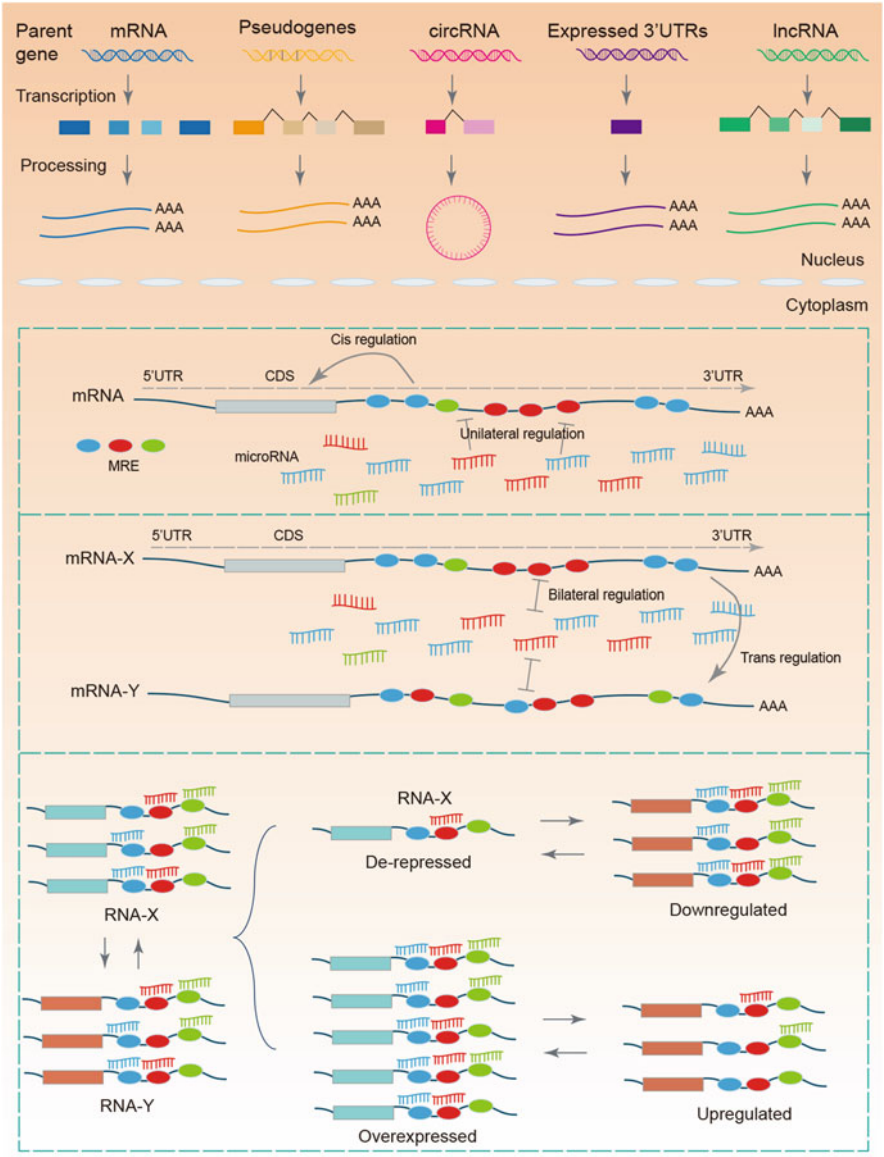


Fig. 6.9 MiRNA-mediated novel RNA regulatory pattern

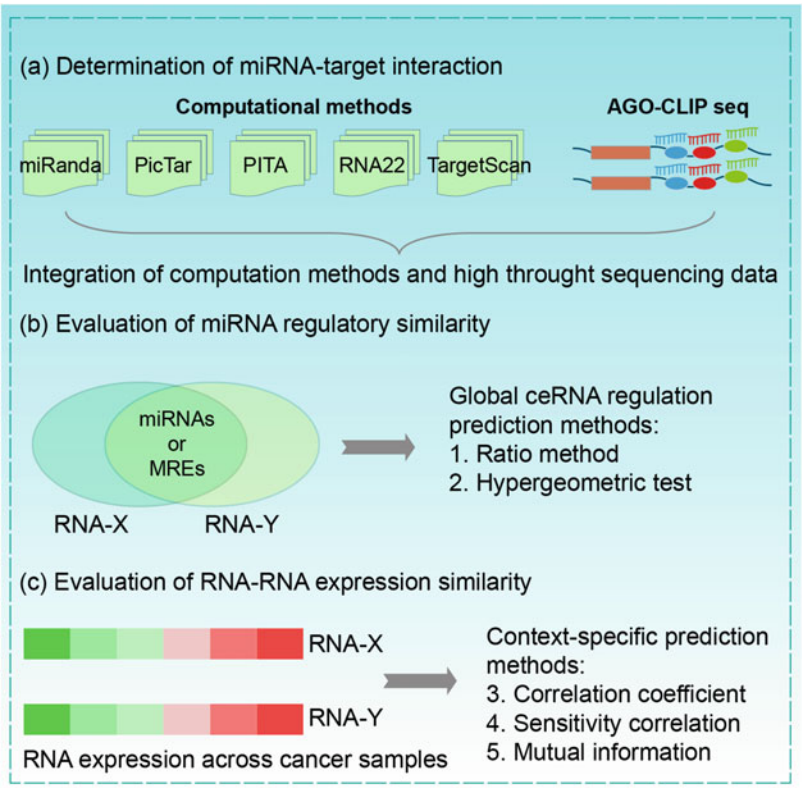


Fig. 6.10 The pipeline used to identify ceRNA regulation

Table 6.2 Summary of computational approaches for identifying miRNA-mediated ceRNA interactions

Methods	Input data	Statistical methods	*P*-value	PMID
Ratio	miRNA–target regulation	No	N	27365046
HyperT	miRNA–target regulation	Hypergeometric test	Y	24297251
HyperC	miRNA–target regulation; RNA expression	Hypergeometric test; correlation coefficient	Y	26304537
SC	miRNA–target regulation; miRNA and RNA expression	Sensitive correlation coefficient; random test	Y	25033876
CMI	miRNA–target regulation; miRNA and RNA expression	CMI; random test	Y	22000015

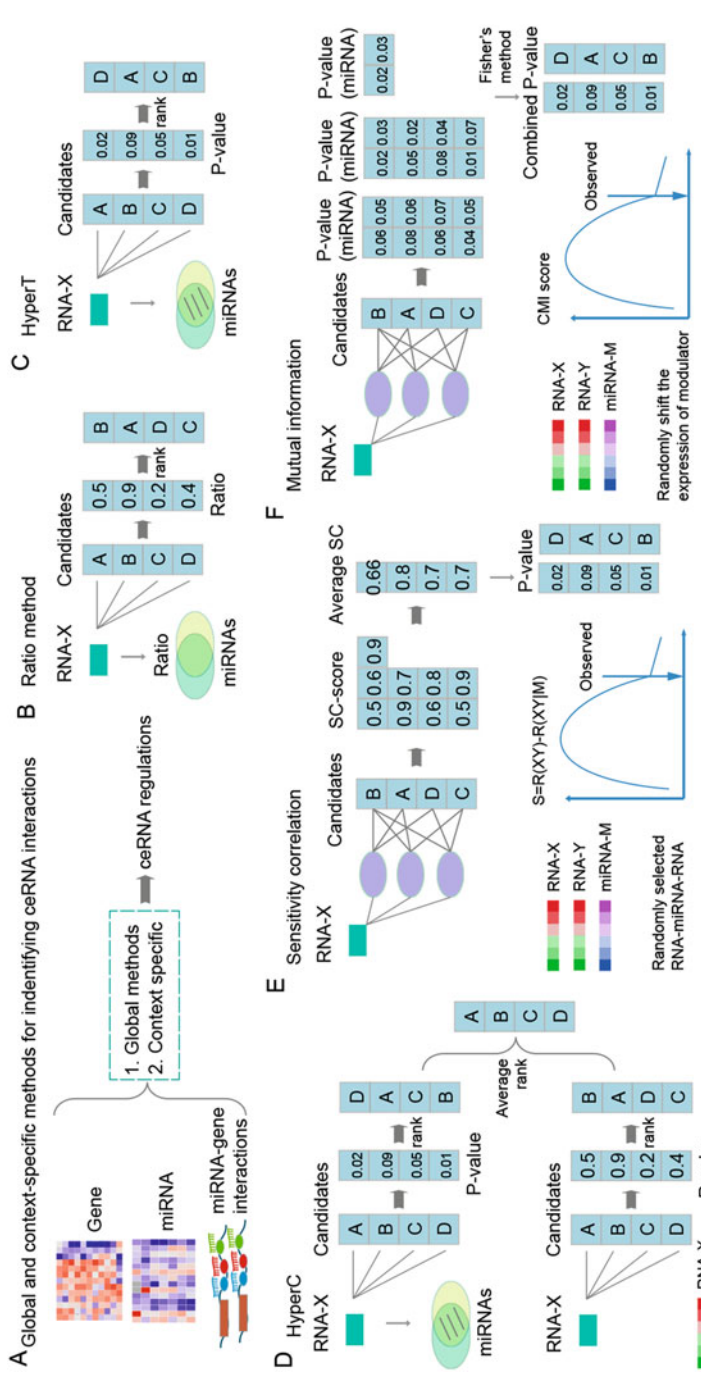


Fig. 6.11 The pipeline of methods to identify miRNA-mediated ceRNA regulations. (**a**) The pipeline for identifying global and context-specific ceRNA regulation. (**b**) The pipeline for the ratio-based method. (**c**) The pipeline for the hypergeometric test-based method, termed HyperT. (**d**) The pipeline for the combination of hypergeometric test and co-expression method, termed HyperC. (**e**) The pipeline for the SC-based method. (**f**) The pipeline for the CMI-based methods

1. *Ratio-based method*. Based on the ceRNA hypothesis that the proportion of miRNAs that RNA-Y shares with RNA-X from all candidate gene sets S is calculated as:

$$R_X = \frac{\text{miRNA}_X \cap \text{miRNA}_Y}{\text{miRNA}_X}, X \in S.$$

where miRNA_X is the miRNA set that regulates RNA-X and miRNA_Y is the miRNA set that regulates RNA-Y (Fig. 6.11b).

2. *Hypergeometric test-based method*. The hypergeometric test is used to determine whether RNA-X and RNA-Y are co-regulated by miRNAs (Fig. 6.11c). This statistic test computes the significance of common miRNAs for each ceRNA pair. The probability P is calculated as follows:

$$P = 1 - F(N_{XY} - 1|N, N_X, N_Y) = 1 - \sum_{t=0}^{N_{XY}} \frac{\binom{N_X}{t}\binom{N-N_X}{N_Y-t}}{\binom{N}{N_Y}}.$$

where N represents the number of all miRNAs in the human genome, N_X and N_Y are the total number of miRNAs that regulate RNA-X and RNA-Y, respectively, and N_{XY} represents the number of miRNAs shared between RNA-X and RNA-Y. P-values are subject to false discovery rate (FDR) correction and RNAs are ranked based on the FDR values.

3. *Hypergeometric test combined with co-expression-based method*. To identify the active ceRNA–ceRNA regulatory pairs in a given context, the co-expression principle is used to filter the ceRNA regulatory interactions that were identified based on the above two global predictions (Chiu et al. 2015b; Zhou et al. 2014). Based on RNA-X and RNA-Y expression profile datasets, the Pearson correlation coefficient (R) of each candidate ceRNA regulation pair discovered is calculated as:

$$R = \frac{\sum_{i=1}^{n}(\text{expr}X_i - \bar{\text{expr}}X)(\text{expr}Y_i - \bar{\text{expr}}Y)}{\sqrt{\sum_{i=1}^{n}(\text{expr}X_i - \bar{\text{expr}}X)^2}\sqrt{\sum_{i=1}^{n}(\text{expr}Y_i - \bar{\text{expr}}Y)^2}}.$$

where $\text{expr}X_i$ and $\text{expr}Y_i$ are the expression levels of RNA-X and RNA-Y in sample i in a specific context, and $\bar{\text{expr}}X$ and $\bar{\text{expr}}Y$ are the average expression levels of RNA-X and RNA-Y across all tumor samples. The RNA-X and RNA-Y are ranked by the P-value of hypergeometric test and the correlation coefficient, separately. Finally, the average rank of each ceRNA regulatory is calculated to rank the candidate genes (Fig. 6.11d).

4. *SC-based method*. In addition, miRNA expression profiles can be also integrated to identify ceRNA regulation in cancer. There is a common method that is used to discover highly correlated ceRNA pairs in which the correlation is due to the presence of one or more miRNAs (Fig. 6.11e). SC had been proposed to identify a sponge interaction network between mRNAs and lncRNAs in breast cancer (Paci et al. 2014). For a candidate pair of RNA-X and RNA-Y, given a co-regulated miRNA-M, the method is as follows:

$$R_{\mathrm{XY|M}} = \frac{R_{\mathrm{XY}} - R_{\mathrm{XM}} R_{\mathrm{MY}}}{\sqrt{1 - R_{\mathrm{XY}}^2}\sqrt{1 - R_{\mathrm{MY}}^2}}.$$

where, R_{XY}, R_{XM}, and R_{MY} represent the Pearson correlation coefficient between RNA-X and RNA-Y, RNA-X and miRNA-M, RNA-Y and miRNA-M, respectively. The SC of miRNA-M, termed S, and the corresponding candidate ceRNA pairs are calculated as:

$$S = R_{\mathrm{XY}} - R_{\mathrm{XY}|M}.$$

Finally, to calculate the significant correlation, a random background distribution of the S was generated by calculating the score S of randomly selected combinations of RNA-X/miRNA/RNA-Y ceRNA regulation.

5. *CMI-based method*. Based on the CMI method (Sumazin et al. 2011), the ceRNA–ceRNA interactions are identified (Fig. 6.11f). The RNA–RNA correlations and miRNAs/RNAs expression profiles are required in the method. We can use Hermes method, which predicts ceRNA interactions from expression profiles of candidate RNAs and their common miRNA regulators using CMI. First, the size of the common miRNAs that regulate two candidate RNAs is necessary to be statistically significant relative to the two individual miRNA sizes, and this is performed by Fisher's exact test. Then Hermes evaluates the statistical significance of the test for each miRNA i as:

$$I(\mathrm{miRNA}_i, A|X) > I(\mathrm{miRNA}_i, A).$$

The random test where the candidate modulator's expressions across cancer patients are perturbed and calculated for P-value of each RNA-X/miRNA/RNA-Y interaction. The final significance for all miRNAs is then calculated by combining all the individual P-values for each miRNA i. This is based on Fisher's method,

$$\chi^2 = -2\sum_{k=1}^{N} \ln(p_k).$$

where N is the total number of miRNAs.

6.5 Future Directions

Computational analysis of miRNA regulatory networks represents a robust platform for understanding miRNA functions in complex diseases, including cancer. Here, we summarized the widely used computational methods for identifying miRNA–gene, miRNA–miRNA, and ceRNA–ceRNA regulation in cancer. However, gene expression is also regulated by other regulatory factors, such as transcription factors and RNA binding proteins. There are both opportunities and challenges for the comprehensive integration of these regulatory layers for understanding gene expression regulation in cancer.

References

Aftabuddin M, Mal C, Deb A, Kundu S (2014) C2Analyzer: co-target-co-function analyzer. Genomics Proteomics Bioinformatics 12(3):133–136. https://doi.org/10.1016/j.gpb.2014.03.003

Agarwal V, Bell GW, Nam JW, Bartel DP (2015) Predicting effective microRNA target sites in mammalian mRNAs. elife 4. https://doi.org/10.7554/eLife.05005

Ala U, Karreth FA, Bosia C, Pagnani A, Taulli R, Léopold V, Tay Y, Provero P, Zecchina R, Pandolfi PP (2013) Integrated transcriptional and competitive endogenous RNA networks are cross-regulated in permissive molecular environments. Proc Natl Acad Sci USA 110(18):7154–7159. https://doi.org/10.1073/pnas.1222509110

Antonov AV, Dietmann S, Wong P, Lutter D, Mewes HW (2009) GeneSet2miRNA: finding the signature of cooperative miRNA activities in the gene lists. Nucleic Acids Res 37(Web Server issue):W323–W328. https://doi.org/10.1093/nar/gkp313

Bajan S, Hutvagner G (2014) Regulation of miRNA processing and miRNA mediated gene repression in cancer. Microrna 3(1):10–17. https://doi.org/10.2174/2211536602666140110234046

Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136(2):215–233. https://doi.org/10.1016/j.cell.2009.01.002

Bartel DP (2018) Metazoan MicroRNAs. Cell 173(1):20–51. https://doi.org/10.1016/j.cell.2018.03.006

Chang TC, Mendell JT (2007) microRNAs in vertebrate physiology and human disease. Annu Rev Genomics Hum Genet 8:215–239. https://doi.org/10.1146/annurev.genom.8.080706.092351

Chaulk SG, Ebhardt HA, Fahlman RP (2016) Correlations of microRNA:microRNA expression patterns reveal insights into microRNA clusters and global microRNA expression patterns. Mol BioSyst 12(1):110–119. https://doi.org/10.1039/c5mb00415b

Chen D, Fu LY, Zhang Z, Li G, Zhang H, Jiang L, Harrison AP, Shanahan HP, Klukas C, Zhang HY, Ruan Y, Chen LL, Chen M (2014) Dissecting the chromatin interactome of microRNA genes. Nucleic Acids Res 42(5):3028–3043. https://doi.org/10.1093/nar/gkt1294

Chiu HS, Llobet-Navas D, Yang X, Chung WJ, Ambesi-Impiombato A, Iyer A, Kim HR, Seviour EG, Luo Z, Sehgal V, Moss T, Lu Y, Ram P, Silva J, Mills GB, Califano A, Sumazin P (2015a) Cupid: simultaneous reconstruction of microRNA-target and ceRNA networks. Genome Res 25(2):257–267. https://doi.org/10.1101/gr.178194.114

Chiu YC, Hsiao TH, Chen Y, Chuang EY (2015b) Parameter optimization for constructing competing endogenous RNA regulatory network in glioblastoma multiforme and other cancers. BMC Genomics 16(Suppl 4):S1. https://doi.org/10.1186/1471-2164-16-s4-s1

da Silveira WA, Renaud L, Simpson J, Glen WB Jr, Hazard ES, Chung D, Hardiman G (2018) miRmapper: a tool for interpretation of miRNA¯mRNA interaction networks. Genes (Basel) 9(9). https://doi.org/10.3390/genes9090458

Dweep H, Gretz N (2015) miRWalk2.0: a comprehensive atlas of microRNA-target interactions. Nat Methods 12(8):697. https://doi.org/10.1038/nmeth.3485

Friedman RC, Farh KK, Burge CB, Bartel DP (2009) Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 19(1):92–105. https://doi.org/10.1101/gr.082701.108

Friedman Y, Naamati G, Linial M (2010) MiRror: a combinatorial analysis web tool for ensembles of microRNAs and their targets. Bioinformatics 26(15):1920–1921. https://doi.org/10.1093/bioinformatics/btq298

Gennarino VA, Sardiello M, Avellino R, Meola N, Maselli V, Anand S, Cutillo L, Ballabio A, Banfi S (2009) MicroRNA target prediction by expression analysis of host genes. Genome Res 19(3):481–490. https://doi.org/10.1101/gr.084129.108

Guo L, Sun B, Wu Q, Yang S, Chen F (2012) miRNA-miRNA interaction implicates for potential mutual regulatory pattern. Gene 511(2):187–194. https://doi.org/10.1016/j.gene.2012.09.066

Hsu CW, Juan HF, Huang HC (2008) Characterization of microRNA-regulated protein-protein interaction network. Proteomics 8(10):1975–1979. https://doi.org/10.1002/pmic.200701004

Hua L, Zhou P, Li L, Liu H, Yang Z (2013) Prioritizing breast cancer subtype related miRNAs using miRNA-mRNA dysregulated relationships extracted from their dual expression profiling. J Theor Biol 331:1–11. https://doi.org/10.1016/j.jtbi.2013.04.008

Hua L, Xia H, Zhou P, Li D, Li L (2014) Combination of microRNA expression profiling with genome-wide SNP genotyping to construct a coronary artery disease-related miRNA-miRNA synergistic network. Biosci Trends 8(6):297–307. https://doi.org/10.5582/bst.2014.01031

Huang Z, Shi J, Gao Y, Cui C, Zhang S, Li J, Zhou Y, Cui Q (2019) HMDD v3.0: a database for experimentally supported human microRNA-disease associations. Nucleic Acids Res 47(D1):D1013–d1017. https://doi.org/10.1093/nar/gky1010

Jia R, Chai P, Zhang H, Fan X (2017) Novel insights into chromosomal conformations in cancer. Mol Cancer 16(1):173. https://doi.org/10.1186/s12943-017-0741-5

Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I, Papadimitriou D, Kavakiotis I, Maniou S, Skoufos G, Vergoulis T, Dalamagas T, Hatzigeorgiou AG (2018) DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. Nucleic Acids Res 46(D1):D239–d245. https://doi.org/10.1093/nar/gkx1141

Kobayashi H, Tomari Y (2016) RISC assembly: coordination between small RNAs and Argonaute proteins. Biochim Biophys Acta 1859(1):71–81. https://doi.org/10.1016/j.bbagrm.2015.08.007

Kozomara A, Birgaoanu M, Griffiths-Jones S (2019) miRBase: from microRNA sequences to function. Nucleic Acids Res 47(D1):D155–d162. https://doi.org/10.1093/nar/gky1141

Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M, Rajewsky N (2005) Combinatorial microRNA target predictions. Nat Genet 37(5):495–500. https://doi.org/10.1038/ng1536

Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of mammalian microRNA targets. Cell 115(7):787–798. https://doi.org/10.1016/s0092-8674(03)01018-3

Li YS, Huo CQ, Lin XY, Xu J (2018) Computational identification of cross-talking ceRNAs. Adv Exp Med Biol 1094:97–108. https://doi.org/10.1007/978-981-13-0719-5_10

Li Y, Jin X, Wang Z, Li L, Chen H, Lin X, Yi S, Zhang Y, Xu J (2019) Systematic review of computational methods for identifying miRNA-mediated RNA-RNA crosstalk. Brief Bioinform 20(4):1193–1204. https://doi.org/10.1093/bib/bbx137

Liang H, Li WH (2007) MicroRNA regulation of human protein protein interaction network. RNA 13(9):1402–1408. https://doi.org/10.1261/rna.634607

Lu TX, Rothenberg ME (2018) MicroRNA. J Allergy Clin Immunol 141(4):1202–1207. https://doi.org/10.1016/j.jaci.2017.08.034

Mal C, Deb A, Aftabuddin M, Kundu S (2015) A network analysis of miRNA mediated gene regulation of rice: crosstalk among biological processes. Mol BioSyst 11(8):2273–2280. https://doi.org/10.1039/c5mb00222b

Meng X, Wang J, Yuan C, Li X, Zhou Y, Hofestädt R, Chen M (2015) CancerNet: a database for decoding multilevel molecular interactions across diverse cancer types. Oncogenesis 4(12): e177. https://doi.org/10.1038/oncsis.2015.40

Paci P, Colombo T, Farina L (2014) Computational analysis identifies a sponge interaction network between long non-coding RNAs and messenger RNAs in human breast cancer. BMC Syst Biol 8:83. https://doi.org/10.1186/1752-0509-8-83

Poliseno L, Salmena L, Zhang J, Carver B, Haveman WJ, Pandolfi PP (2010) A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. Nature 465(7301): 1033–1038. https://doi.org/10.1038/nature09144

Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP (2011) A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell 146(3):353–358. https://doi.org/10.1016/j.cell.2011.07.014

Segura J, Sorzano CO, Cuenca-Alba J, Aloy P, Carazo JM (2015) Using neighborhood cohesiveness to infer interactions between protein domains. Bioinformatics 31(15):2545–2552. https://doi.org/10.1093/bioinformatics/btv188

Shalgi R, Lieber D, Oren M, Pilpel Y (2007) Global and local architecture of the mammalian microRNA-transcription factor regulatory network. PLoS Comput Biol 3(7):e131. https://doi.org/10.1371/journal.pcbi.0030131

Shao T, Wang G, Chen H, Xie Y, Jin X, Bai J, Xu J, Li X, Huang J, Jin Y, Li Y (2019) Survey of miRNA-miRNA cooperative regulation principles across cancer types. Brief Bioinform 20(5): 1621–1638. https://doi.org/10.1093/bib/bby038

Sumazin P, Yang X, Chiu HS, Chung WJ, Iyer A, Llobet-Navas D, Rajbhandari P, Bansal M, Guarnieri P, Silva J, Califano A (2011) An extensive microRNA-mediated network of RNA-RNA interactions regulates established oncogenic pathways in glioblastoma. Cell 147(2):370–381. https://doi.org/10.1016/j.cell.2011.09.041

Sun J, Zhou M, Yang H, Deng J, Wang L, Wang Q (2013) Inferring potential microRNA-microRNA associations based on targeting propensity and connectivity in the context of protein interaction network. PLoS One 8(7):e69719. https://doi.org/10.1371/journal.pone.0069719

Teng L, He B, Wang J, Tan K (2015) 4DGenome: a comprehensive database of chromatin interactions. Bioinformatics 31(15):2560–2564. https://doi.org/10.1093/bioinformatics/btv158

Vergoulis T, Vlachos IS, Alexiou P, Georgakilas G, Maragkakis M, Reczko M, Gerangelos S, Koziris N, Dalamagas T, Hatzigeorgiou AG (2012) TarBase 6.0: capturing the exponential growth of miRNA targets with experimental support. Nucleic Acids Res 40 (Database issue): D222–229. https://doi.org/10.1093/nar/gkr1161

Vlachos IS, Kostoulas N, Vergoulis T, Georgakilas G, Reczko M, Maragkakis M, Paraskevopoulou MD, Prionidis K, Dalamagas T, Hatzigeorgiou AG (2012) DIANA miRPath v.2.0: investigating the combinatorial effect of microRNAs in pathways. Nucleic Acids Res 40 (Web Server issue): W498–504. https://doi.org/10.1093/nar/gks494

Vlachos IS, Paraskevopoulou MD, Karagkouni D, Georgakilas G, Vergoulis T, Kanellos I, Anastasopoulos IL, Maniou S, Karathanou K, Kalfakakou D, Fevgas A, Dalamagas T, Hatzigeorgiou AG (2015a) DIANA-TarBase v7.0: indexing more than half a million experimentally supported miRNA:mRNA interactions. Nucleic Acids Res 43(Database issue): D153–159. https://doi.org/10.1093/nar/gku1215

Vlachos IS, Zagganas K, Paraskevopoulou MD, Georgakilas G, Karagkouni D, Vergoulis T, Dalamagas T, Hatzigeorgiou AG (2015b) DIANA-miRPath v3.0: deciphering microRNA function with experimental support. Nucleic Acids Res 43(W1):W460–W466. https://doi.org/10.1093/nar/gkv403

Xie X, Ma W, Songyang Z, Luo Z, Huang J, Dai Z, Xiong Y (2016) CCSI: a database providing chromatin-chromatin spatial interaction information. Database (Oxford) 2016. https://doi.org/10.1093/database/bav124

Xu J, Li CX, Li YS, Lv JY, Ma Y, Shao TT, Xu LD, Wang YY, Du L, Zhang YP, Jiang W, Li CQ, Xiao Y, Li X (2011) MiRNA-miRNA synergistic network: construction via co-regulating functional modules and disease miRNA topological features. Nucleic Acids Res 39(3): 825–836. https://doi.org/10.1093/nar/gkq832

Xu J, Li YS, Lu JP, Pan T, Ding N, Wang ZS, Shao TT, Zhang JW, Wang LH, Li X (2015) The mRNA related ceRNA-ceRNA landscape and significance across 20 major cancer types. Nucleic Acids Res 43(17):8169–8182. https://doi.org/10.1093/nar/gkv853

Xu J, Feng L, Han ZJ, Li YS, Wu AW, Shao TT, Ding N, Li LL, Deng W, Di XB, Wang J, Zhang LF, Li X, Zhang KT, Cheng SJ (2016) Extensive ceRNA-ceRNA interaction networks mediated by miRNAs regulate development in multiple rhesus tissues. Nucleic Acids Res 44(19): 9438–9451. https://doi.org/10.1093/nar/gkw587

Xu J, Shao T, Ding N, Li Y, Li X (2017) miRNA-miRNA crosstalk: from genomics to phenomics. Brief Bioinform 18(6):1002–1011. https://doi.org/10.1093/bib/bbw073

Xu J, Shao TT, Song MX, Xie YJ, Zhou JL, Yin JQ, Ding N, Zou HZ, Li YS, Zhang JW (2020) MIR22HG acts as a tumor suppressor via TGF beta/SMAD signaling and facilitates immunotherapy in colorectal cancer. Mol Cancer 19(1). https://doi.org/10.1186/S12943-020-01174-W

Zhang C (2008) MicroRNomics: a newly emerging approach for disease biology. Physiol Genomics 33(2):139–147. https://doi.org/10.1152/physiolgenomics.00034.2008

Zhou X, Liu J, Wang W (2014) Construction and investigation of breast-cancer-specific ceRNA network based on the mRNA and miRNA expression data. IET Syst Biol 8(3):96–103. https://doi.org/10.1049/iet-syb.2013.0025

Chapter 7
Analysis of the p53/microRNA Network in Cancer

Markus Kaller, Sabine Hünten, Helge Siemens, and Heiko Hermeking

Abstract MicroRNAs (miRNAs) are important components of the signaling cascades that mediate and regulate tumor suppression exerted by p53. This review illustrates some of the main principles that underlie the mechanisms by which miRNAs participate in p53's function and how they were identified. Furthermore, the current status of the research on the connection between p53 and miRNAs, as well as alterations in the p53/miRNA pathways found in cancer will be summarized and discussed. In addition, experimental and bioinformatic approaches which can be applied to study the connection between p53 and miRNAs are described. Although, some of the central miRNA-encoding genes that mediate the effects of p53, such as the *miR-34* and *miR-200* families, have been identified, much more analyses remain to be performed to fully elucidate the connections between p53 and miRNAs.

Keywords p53 · microRNA · miRNA · Tumor suppression · SILAC · Next generation sequencing · Genome-wide analysis · miR-34 · miR-34a · miR-34b/c

---

Analysis of the p53/microRNA network in cancer" by Kaller, Hünten, Siemens and Hermeking), please include the footnote to the abstract" This chapter is an update of the previously published work: Hünten S., Siemens H., Kaller M., Hermeking H. (2013) The p53/microRNA Network in Cancer: Experimental and Bioinformatics Approaches. In: Schmitz U., Wolkenhauer O., Vera J. (eds) MicroRNA Cancer Regulation

Markus Kaller, Sabine Hünten and Helge Siemens contributed equally with all other contributors.

M. Kaller · S. Hünten · H. Siemens
Experimental and Molecular Pathology, Institute of Pathology, Ludwig-Maximilians-University, Munich, Germany

H. Hermeking (✉)
Experimental and Molecular Pathology, Institute of Pathology, Ludwig-Maximilians-University, Munich, Germany

German Cancer Consortium (DKTK), Partner site Munich, Munich, Germany

German Cancer Research Center (DKFZ), Heidelberg, Germany
e-mail: Heiko.Hermeking@med.uni-muenchen.de

U. Schmitz et al. (eds.), *Systems Biology of MicroRNAs in Cancer*, Advances in Experimental Medicine and Biology 1385,
https://doi.org/10.1007/978-3-031-08356-3_7

7.1 Introduction to p53 Biology

The p53 transcription factor is encoded by a tumor suppressor gene, which is presumably the most commonly mutated gene in human cancer (Soussi 2011). In addition, many of the cancers without *p53* mutation may harbor alterations up- or downstream of p53, which also impede the ability of p53 to suppress tumor cell growth. p53 and its loss may represent attractive targets for tumor therapy (Cheok et al. 2011). Most *p53* mutations target the DNA binding properties of p53, suggesting that the regulation of specific target genes is central to the tumor suppression mediated by p53. However, alternative functions of p53 in the cytoplasm and in mitochondria have also been described (Green and Kroemer 2009). p53's transcriptional activity is induced by various forms of cellular stress that cause diverse posttranslational modifications of p53, which are thought to allow a fine-tuning of the cellular response to the type and extent of stress experienced by the respective cell (Kruse and Gu 2009), summarized in Fig. 7.1). For example, repairable DNA damage may cause a transient cell cycle arrest, whereas extensive damage may induce apoptosis via generating different levels of p53 activity. DNA damage in the form of double-strand DNA breaks was one of the first inducers of p53 to be discovered. Subsequently, ribosomal, replication, metabolic, oxidative, and transcriptional stress, as well as hypoxia were found to cause an increase in p53's transcriptional activity. These alterations stimulate distinct signaling cascades which activate enzymes that modify p53 or regulate cofactors binding to p53. For example, DNA double-strand breaks lead to activation of the ATM kinase, which phosphorylates p53 at multiple N-terminal residues (Derheimer and Kastan 2010) and thereby increases its transactivation activity. Furthermore, p53 may be activated by inhibition of the MDM2 protein, which represents an E3-ubiquitin ligase that marks p53 for proteasomal degradation. p53 forms tetramers that bind to palindromic recognition sites often organized in tandem repeats with spacers of varying lengths between them (Fig. 7.1). Promoters display gradual responsiveness to p53 either due to different numbers of p53 binding motifs or due to the presence of high affinity versus low-affinity sites (Menendez et al. 2009). For example, the *p21* gene has a high-affinity p53 binding site and mediates cell cycle arrest, whereas genes that mediate cell death harbor low-affinity p53 binding sites. Therefore, apoptosis is presumably only induced when p53 is strongly activated. p53 directly activates a large set of genes, which mediate numerous cellular functions that contribute to tumor suppression (Kastenhuber and Lowe 2017; Fischer 2017). Many, but not all of these protein-coding target genes are depicted in Fig. 7.1. The activation of p53 target genes is either caused by an increase in p53 abundance after p53 stabilization, anti-repression of specific genes after removal of repressive MDM2/MDMX from p53 by acetylation and/or phosphorylation, or formation of promoter-specific transcriptional complexes (Kruse and Gu 2009). Furthermore, p53 may mediate the repression of specific genes. However, the mechanisms of transcriptional repression by p53 are less well understood and are presumably mediated indirectly via the p21/DREAM complex (Fischer et al. 2014, 2016). Furthermore, miRNAs represent important mediators of gene repression caused by p53 (Hermeking 2012).

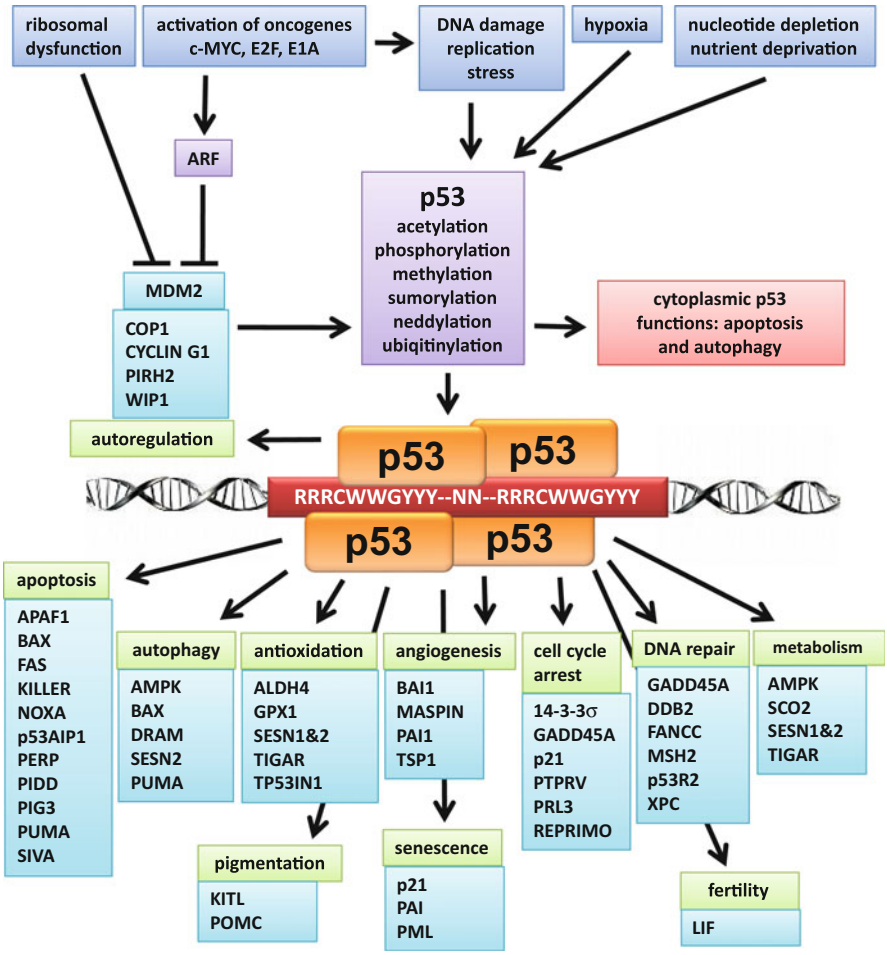


Fig. 7.1 p53 as a central mediator of stress responses. In this model, the types of stress and events leading to activation of p53 and the protein-encoding genes activated by p53 are depicted. p53 is shown as a symbolic tetramer occupying a p53 binding motif (in red) containing two palindromic DNA sequences (white letters with R = purines (A or G), Y = pyrimidines (C or T), W = A or T, and N = bases representing spacers between the 2 palindromic half-sites). Processes regulated by p53 and the respective p53 target genes implicated are indicated

7.2 p53 and the miRNA World: Current State of the Art

miRNAs have presumably evolved to allow organisms to effectively deal with stress (Leung and Sharp 2007, 2010). In line with this notion, the p53 stress-response pathway is heavily interconnected with miRNAs not only by regulating their expression and processing, but also since p53 itself represents a downstream target of miRNAs (see Figs. 7.2, 7.3, and 7.4). The protein-coding genes regulated by p53

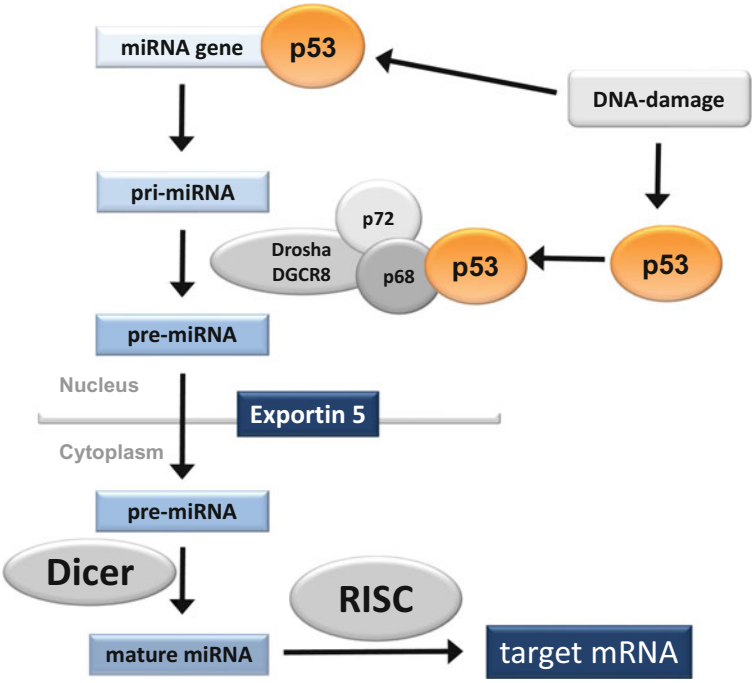


Fig. 7.2 Effects of p53 on the miRNA processing pathway. The synthesis of miRNAs in mammalian cells and the known modes of regulation by p53 are depicted

elicit several cellular phenotypes/processes, which contribute to tumor suppression, for example, induction of cell cycle arrest, senescence, and apoptosis, as well as inhibition of metastasis, angiogenesis, and glycolysis (Vogelstein et al. 2000; Hermeking 2007, 2003, 2010; Vousden and Ryan 2009; Vousden and Prives 2009; Riley et al. 2008). Interestingly, these processes are also regulated and, in some cases, induced by p53-regulated miRNAs (Hermeking 2007, 2010, 2012). Consequently, the characterization of a number of miRNAs directly regulated by p53 and the cellular effects of these connections have been reported. For an overview, see Fig. 7.3.

7.2.1 The miR-34 *Genes*

In 2007, the miR-34 genes, *miR-34a*, and *miR-34b/c*, were reported to be directly regulated by p53 by a number of laboratories using diverse approaches (Tarasov et al. 2007; Chang et al. 2007; He et al. 2007a; Bommer et al. 2007; Raver-Shapira et al. 2007; Corney et al. 2007). For example, we determined the abundance of miRNAs in libraries representing small RNAs generated after p53 activation using a next generation sequencing approach (Tarasov et al. 2007): we found that *miR-34a*

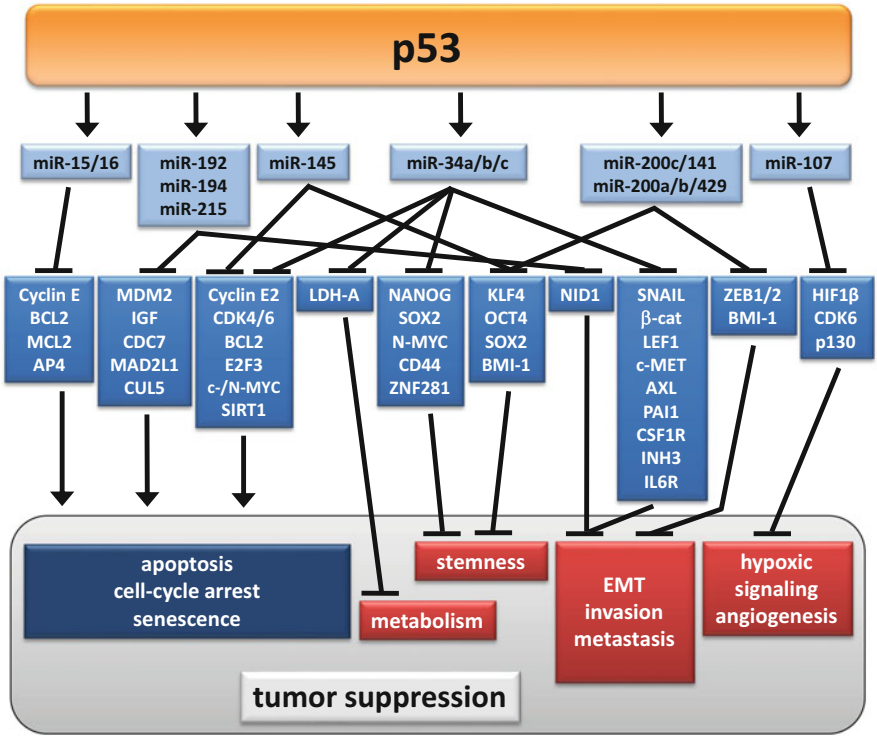


Fig. 7.3 Regulation of miRNA expression by p53. Model summarizing direct transcriptional activation of miRNA-encoding genes, the affected miRNA targets and the reported cellular effects, which collectively contribute to tumor suppression by p53

showed the most pronounced increase among all detected miRNAs after p53 activation, which is mediated by p53 binding sites in the promoter region of its host gene. When ectopically expressed, *miR-34a* and *miR-34b/c* displayed tumor-suppressive activities, i.e., they caused induction of apoptosis and senescence, inhibition of cell cycle progression, and decrease of angiogenesis (reviewed in (Hermeking 2007, 2010; He et al. 2007b). These effects were mediated by direct downregulation of numerous key regulators and effectors of these processes such as BCL-2, Cyclin E, CDK4, and CDK6. Meanwhile, a large number of additional miR-34 targets have been identified using a variety of approaches (reviewed in Hermeking (2010, 2012), Rokavec et al. (2014a); see also Kaller et al. (2011), Lal et al. (2011) and references therein). Among the miR-34 targets, SIRT1, c-MET, AXL, c-/N-MYC, LDHA, and SNAIL seem to be especially relevant for the suppression of cancer. In fact, their common upregulation in tumors could be due to the frequent inactivation of the p53/miR-34 axis during tumor development (Lodygin et al. 2008; Vogt et al. 2011). These targets contribute to the suppression of migration and invasion (SNAIL, c-MET, AXL) and metabolism (LDHA). In the case of c-MET it was shown that p53 downregulates c-*MET* expression via

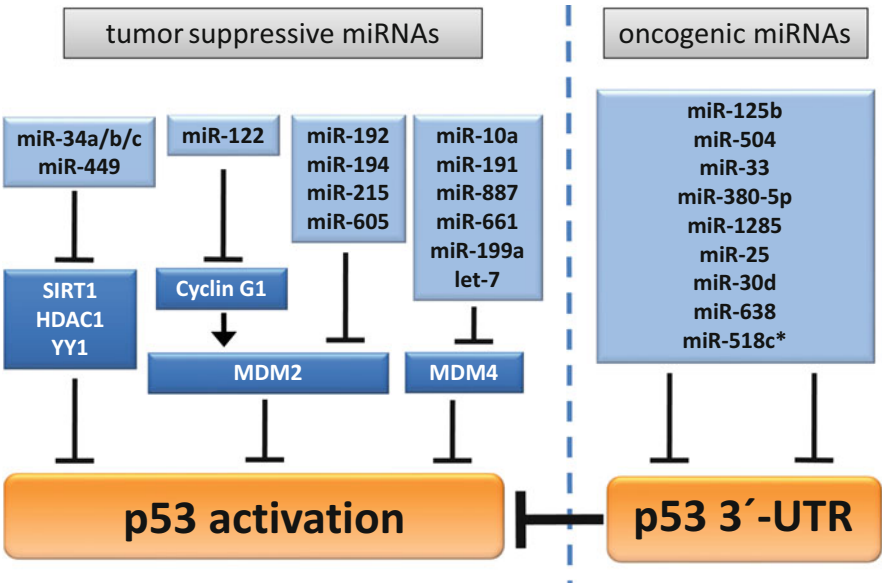


Fig. 7.4 Regulation of p53 by miRNAs. Model summarizing the regulation of the 3'-UTR of p53, as well as the downregulation of p53-modifying enzymes by miRNAs. Oncogenic miRNAs with seed-matching sequences in the 3'-UTR of p53 or tumor-suppressive miRNAs directly downregulating the indicated proteins are depicted

SP1-mediated occupancy and repression of the c-*MET* promoter and by inducing miR-34a/b/c, which directly target the 3'-UTR of the c-*MET* mRNA (Hwang et al. 2011). p53 may suppress metastasis by preventing epithelial–mesenchymal transitions, which have been implicated in the early, invasive stages of metastasis. Accordingly, p53 activation promotes mesenchymal–epithelial transition (MET) and favors the epithelial state of cells (Schubert and Brabletz 2011). We found that p53-induced MET is mediated by the induction of *miR-34a* and *miR-34b/c* in colorectal cancer cell lines (Siemens et al. 2011). miR-34a and miR-34b/c achieve this effect by negatively regulating a master-regulator of EMT, the SNAIL transcription factor (Siemens et al. 2011; Kim et al. 2011b). In addition, we found that the *miR-34a* and the *miR-34b/c* genes are directly repressed by SNAIL (Siemens et al. 2011). Therefore, miR-34a/b/c and SNAIL form a double-negative feedback loop (summarized in Brabletz (2012)). Stemness represents another important oncogenic trait of cancer cells which is suppressed by miR-34. It was shown that miR-34 directly suppresses CD44, which blocks the expansion of cancer-initiating tumor stem cells in a mouse model of prostate cancer (Liu et al. 2011). When *miR-34a* is ectopically expressed, stemness markers such as CD133, CD44, and BMI-1 are downregulated in colorectal cancer cells (Siemens et al. 2011). Furthermore, it was reported that similar to p53, the miR-34 miRNAs provide a barrier for somatic cell reprograming and the generation of IPS (induced pluripotent stem) cells from mouse embryo fibroblast (Choi et al. 2011). miR-34 mediated this effect by downregulation

of NANOG, SOX2, and N-MYC. Therefore, cancer cells with loss of miR-34 expression may also be more prone to become tumor-initiating cells, which exhibit features of stem cells. Furthermore, miR-34 inhibits components of the WNT/β-catenin/TCF pathway, such as β-catenin, LEF1, and WNT1 (Kaller et al. 2011; Kim et al. 2011a). Thereby, miR-34 may contribute to the suppression of stemness- and EMT-related features of cancer cells. In the recent years, we identified and characterized additional miR-34 targets, which are critical for the suppression of stemness and EMT by miR-34 in cancer cells, such as ZNF281 (Hahn et al. 2013), CSF1R (Shi et al. 2020), INH3/PPP1R11 (Li et al. 2017), IL6R (Rokavec et al. 2014b) and PAI1 (Oner et al. 2018).

The miR-34 family also includes miR-449. Although the seed sequences of miR-34a/b/c and miR-449a/b/c are highly conserved, the regulation of the genes encoding these miRNAs is divergent as miR-449 expression is induced by E2F1, but not by p53 and/or DNA damage (Lize et al. 2011). Therefore, the regulation of similar targets by miR-34 and miR-449 miRNAs may occur under rather distinct circumstances. Furthermore, miR-449 presumably has a restricted expression pattern, since it was found to be highly expressed in differentiating lung epithelia and at comparatively low levels in other tissues (Lize et al. 2011).

7.2.2 The miR-200 Family

The two genes encoding the miR-200 family, which give rise to the miR-200c/141 and the miR-200a/200b/429 miRNAs, were identified as direct p53 target genes that enforce mesenchymal–epithelial transitions (MET) (Chang et al. 2011; Kim et al. 2011c) by targeting the EMT regulators ZEB1 and ZEB2 (Gregory et al. 2008). In addition, miR-200 downregulates KLF4 and the polycomb repressor BMI-1, both stemness factors, and thereby contributes to the loss of metastatic capacity of tumor-initiating cancer stem cells (Chang et al. 2011). Therefore, the induction of the miR-200 family represents a new mechanism by which p53 suppresses metastasis (reviewed in Keck and Brabletz 2011; Schubert and Brabletz 2011).

7.2.3 The miR-192 Family

The 3 members of the miRNA-192 family were found to be encoded by p53 target genes using a microarray analysis to monitor miRNA expression after treatment with the Mdm-2 inhibitor Nutlin-3a (Braun et al. 2008). These authors also found that ectopic miR-192 expression induces p21 in a p53-dependent manner. Later it was shown that the miR-192 family targets the IGF pathway and also Mdm-2, which results in the activation of p53 (Pichiorri et al. 2010). Furthermore, this tumor-suppressive loop is impaired in multiple myeloma, which shows downregulation of the miR-192 family. In addition, ectopic miR-192 leads to a G_1 and G_2/M cell cycle

arrest by targeting CDC7, MAD2L1, and CUL5 (Georges et al. 2008). More recently, we showed that p53 inhibits the paracrine induction of EMT by secreted Nidogen-1 (NID1) via direct inducing miR-192 and miR-215, which directly target NID1 in colorectal cancer cells (Rokavec et al. 2019).

7.2.4 Additional p53-Regulated miRNAs

miR-107 is encoded by an intron of the p53-induced *PANK1* gene (Yamakuchi et al. 2010). Ectopic expression of miR-107 decreases HIF1β expression which diminishes the response to hypoxia and blocks tumor angiogenesis and growth. In addition, miR-107 targets the cell cycle regulators CDK6 and p130/pRBL2 (Bohlig et al. 2011).

miR-145 represents a p53-inducible miRNA, which was shown to contribute to the repression of c-*MYC* by p53 via directly targeting the c-*MYC* 3'-UTR (Sachdeva et al. 2009). Interestingly, miR-145 negatively regulates OCT4, SOX2, and KLF4, and thereby represses pluripotency in human embryonic stem cells (Xu et al. 2009). Therefore, miR-145 may at least in part explain why deletion of p53 strongly enhances the generation of IPS cells and potentially promotes the expansion of cancer stem cells (Krizhanovsky and Lowe 2009).

miR-15a and miR-16-1 are encoded by an intron of the *DLEU2* long noncoding (lnc) RNA. Initially, miR-15/16 was shown to be processed at an increased rate after p53 activation (Suzuki et al. 2009; Tarasov et al. 2007). Later, the *DLEU2* gene was shown to be a transcriptional target of p53 (Fabbri et al. 2011; Shi et al. 2014). Since miR-15/16 targets BCL2 and Cyclin E, they affect both, the cell cycle and apoptosis. Moreover, we showed that direct targeting of AP4 by miR-15/16 results in mesenchymal-epithelial transition (MET), inhibition of migration and invasion, and suppression of metastasis of colorectal cancer cells (Shi et al. 2014). Recently, we showed that down-regulation of AP4 also results in DNA damage, chromosomal instability (CIN), and cellular senescence (Chou et al. 2022).

More recently, miR-30e-5p has been reported to be a direct transcriptional target of p53 which inhibits colorectal cancer invasion and metastasis by targeting ITGA6 and ITGB1 (Laudato et al. 2017). Moreover, miR-30a is directly activated by p53 and controls tumor cell invasion and distal spreading via targeting ZEB2 in triple-negative breast cancer (di Gennaro et al. 2018).

The host genes *C9ORF3* and *COL27A1* of the intronic miR-27b-3p and miR-455-3p miRNAs are directly activated by p53, and promote cancer cell quiescence via targeting CKS1B and CAC1, ultimately leading to stabilization of p27 due to its reduced polyubiquitination (La et al. 2018).

7.2.5 *Direct Regulation of p53 Expression by miRNAs*

Several publications demonstrated that miRNAs contribute to the tight control under which p53 is placed in the cell by directly interacting with the 3'-UTR of p53 (reviewed in Liu et al. 2017 and summarized in Fig. 7.4). By computational analysis of putative miRNA binding sites using TargetScan and mirBase prediction software, a binding site of miR-125b was identified in the 3′-UTR of p53 (Careccia et al. 2009). MiR-125b is expressed at high levels in the brain and conserved between human, zebrafish, and other vertebrates. Ectopic expression of miR-125b decreased p53 protein levels and apoptosis in human cells, whereas inhibition of miR-125b had the opposite effect in lung fibroblasts and zebrafish brain. When zebrafishes were treated with DNA damaging agents, miR-125b expression was downregulated, presumably allowing the observed increase in p53 protein. Analysis of 89 colorectal cancer samples revealed that elevated expression of miR-125b is associated with increased tumor size and invasion, and also correlates with poor prognosis and decreased survival (Nishida et al. 2011). These results are in accordance with negative regulation of p53 by miR-125b.

By an in silico search, two miR-504 seed-matching sequences were identified in the 3'-UTR of p53 (Hu et al. 2010). Accordingly, ectopic expression of miR-504 downregulated p53 protein levels, reduced p53-dependent apoptosis and cell cycle arrest, and resulted in increased in vivo tumor formation.

miR-33 also targets p53 by binding to two seed-matching motifs in the 3'-UTR of p53 (Herrera-Merchan et al. 2010). Interestingly, miR-33 is downregulated in hematopoietic stem cells (HSC) and upregulated in more differentiated progenitor cells in so-called super-p53 mice, which are endowed with an extra copy of p53. Ectopic expression of miR-33 in HSC results in increased stemness and decreased recipient survival. In mouse embryonic fibroblasts miR-33 promotes neoplastic transformation presumably via downregulation of p53.

miR-380-5p was found to downregulate p53 in neuroblastomas, which commonly express wild-type p53 (Swarbrick et al. 2010). Neuroblastomas with elevated expression of miR-380-5p showed a decreased patient survival. Furthermore, miR-380-5p was highly expressed in mouse embryonic stem cells and its ectopic expression cooperated with HRAS in transformation, abrogation of oncogene-induced senescence, and promoted tumor formation in mice. Finally, in vivo delivery of a miR-380-5p antagonist decreased tumor size in an orthotopic mouse model of neuroblastoma.

In a systematic, bioinformatic screen 107 potential p53-targeting miRNAs were identified using TargetScan (Tian et al. 2010). When these candidates were experimentally tested in a dual-reporter assay, miR-1285 turned out to be the most effective repressor of p53's 3'-UTR activity. In line with these results, miR-1285 decreased p53 mRNA and protein levels by directly binding to the 3'-UTR of p53 via two seed-matching sequences.

In a similar bioinformatic screen using less stringent criteria and 4 different miRNA target prediction methods (Miranda, TargetScan, PicTar, and RNA22)

67 candidate miRNAs with the potential to directly inhibit p53 expression were identified (Kumar et al. 2011). In a subsequent experimental screen only 8 of these had an inhibitory effect on p53-mediated transactivation. Of these, only miR-200a, −30d and − 25 were effective in a dual-reporter assay employing the *p53* 3'-UTR. By mutation of the respective corresponding seed-matching sequences in reporter constructs, only miR-30d and miR-25 were validated as direct regulators of the *p53* 3'-UTR. In contrast, miR-200a presumably affects the *p53* 3'-UTR by indirect regulation, e.g., via modulation of transcription factors that regulate miRNAs, which directly target p53. In a cellular assay, ectopic miR-30d and miR-25 decreased p53 levels, p53 target expression, and downstream effects of p53 as apoptosis, cell cycle arrest, and senescence. The opposite was observed when both miRNAs were inhibited by antagomirs. In line with these observations, miR-25 and miR-30d were found to be upregulated in multiple myeloma cells, which showed a concomitant downregulation of p53 mRNA expression. Furthermore, inhibition of miR-25 and miR-30d induced p53 and apoptosis in a multiple myeloma cell line. Therefore, miR-25 and miR-30d presumably represent oncogenic miRNAs.

More recently, miR-638 and miR-518c* were shown to directly target p53, together with other tumor suppressors such as PTEN and BRCA1. Furthermore, their overexpression increases cell migration, invasion, and proliferation, suggesting an oncogenic role for these miRNAs (Tay et al. 2014).

7.2.6 Indirect Regulation of p53 by miRNAs

Several examples of p53 being subject to indirect regulation by miRNAs via downregulation of upstream regulators of p53 have been documented (Liu et al. 2017). One of the first cases was the regulation of SIRT1 by miR-34a (Yamakuchi et al. 2008). An in silico search for miR-34a targets, which might affect p53 resulted in the analysis and experimental confirmation of SIRT1 as a miR-34a target. As a consequence of SIRT1 downregulation by miR-34a an increase in p53 activity and enhanced expression of its targets such as p21 and PUMA, as well as increased apoptosis was observed. Since miR-34a itself is induced by p53, the regulations connecting miR-34a, SIRT1, and p53 constitute a positive feedback loop. In tumors, this self-activating loop may be disrupted by the silencing of *miR-34* genes by CpG methylation and mutation/inactivation of p53 (Hermeking 2010; Lodygin et al. 2008; Vogt et al. 2011).

Moreover, miR-34a directly regulates MTA2, HDAC1, and YY1, which are additional factors involved in the regulation of p53 protein levels either via deacetylation or ubiquitination (Chen et al. 2011; Zhao et al. 2013; Kaller et al. 2011).

As mentioned above, miR-449 is similar to miR-34, but regulated by other factors like E2F1. When miR-449 was expressed ectopically, it also indirectly activated p53 via directly suppressing the expression of HDAC1 and SIRT1 (Bou Kheir et al. 2011). This may allow additional pathways to increase p53 activity.

Also, miR-122 leads to an upregulation of p53 (Fornari et al. 2009). However, this is achieved even more indirectly, since the downregulation of the miR-122 target Cyclin G1 presumably leads to decreased recruitment of PP2A phosphatase to MDM2, which results in a decreased MDM2 activity and increased p53 levels/activity. In line with this scenario, ectopic miR-122 expression increased the sensitivity of hepatocellular carcinoma-derived cell lines to doxorubicin.

More recently, miR-885-5p was shown to activate p53 and the expression of p53 target genes (Afanasyeva et al. 2011). Although miR-885-5p was shown to target CDK2 and MCM5, the mechanism for the effect on p53 remained unclear.

miR-192/194/215 are transcriptionally induced by p53 and negatively modulate Mdm2 activity (Pichiorri et al. 2010). Interestingly, their ectopic expression enhanced the therapeutic effectiveness of MDM2 inhibitors against multiple myeloma (MM), an incurable B cell neoplasm, in experimental settings. A similar feedback loop was recently described for miR-605, which is also induced by p53 and negatively regulates MDM2 expression (Xiao et al. 2011).

In addition, several miRNAs such as miR-10a, miR-191-5p, miR-887, miR-661, miR-34a, miR-199a-3p, and let-7 repress MDM4, another important regulator of p53 protein levels (Hoffman et al. 2014).

7.2.7 Direct Involvement of p53 in miRNA Processing and Maturation

Since the levels of certain processed miRNAs were increased after p53 activation even in the absence of an induction of the corresponding primary miRNAs (pri-mRNAs), the possibility that p53 may directly affect the processing of miRNAs was analyzed (Suzuki et al. 2009). Indeed, these authors found that p53 interacts with the miRNA processing complex DROSHA through association with the DEAD-box RNA helicase p68 (indicated in Fig. 7.2). Thereby, p53 enhances the processing of specific pri-miRNAs with growth-suppressive function (e.g., miR-16-1, miR-143, and miR-145) to precursor miRNAs (pre-miRNAs) resulting in a significant increase in the corresponding miRNAs. Therefore, direct transcriptional regulation of any miRNA-encoding gene by p53 should not be deduced from the detection of an increase in mature miRNA levels by techniques like miRNA-Seq. Such analysis should be complemented by quantifications of the pri-miRNA levels and detection of p53 occupancy at the promoter of the respective pri-miRNA encoding gene.

Another link between p53 and miRNA processing has been observed in conditional DICER knockout mice (Mudhasani et al. 2008). DICER deficiency and therefore incomplete miRNA maturation induce p53 and p19/ARF, which leads to reduced proliferation and premature senescence. Interestingly, deletion of *Ink4/Arf* or *p53* prevents premature senescence induced by deletion of DICER. Therefore, a p53-dependent checkpoint seems to monitor proper miRNA processing.

7.2.8 *The p53 Relatives p63 and p73 in the Regulation of miRNAs*

The p53 family members p63 and p73 have also been implicated in the regulation of miRNA expression and processing. TAp63 was shown to coordinately regulate DICER and miR-130b to suppress metastasis (Su et al. 2010). In contrast to *p53,* the *p63* and *p73* genes are not affected by mutations in tumors. p73 promotes genome stability and mediates chemosensitivity, whereas p63 largely lacks these p53-like functions and instead promotes proliferation and cell survival. p63 and p73 were shown to be connected via miRNA regulations: p63 represses the expression of miR-193-5p, which targets p73, thereby causing an increase in p73 expression, whereas p73 induces miR-193-5p (Ory and Ellisen 2011). Interestingly, therapeutic inhibition of miR-193-5p effectively blocked tumor progression in an orthotopic tumor model when combined with otherwise ineffective chemotherapy.

7.3 Alterations of the p53/miRNA Network in Human Cancer

Similar to protein-coding genes, miRNA-encoding genes may harbor oncogenic or tumor-suppressive functions. As discussed above, p53-induced miRNAs promote tumor-suppressive processes like cell cycle arrest, senescence, inhibition of EMT, and metastasis. During cancer initiation or progression, cells with inactivation of miRNA-encoding genes may have a selective advantage, since they presumably display a weakened or missing induction of these tumor suppressive mechanisms. In tumors, miRNA-encoding genes may be inactivated by a number of different mechanisms. The p53-inducible miRNAs discussed above are likely to be downregulated in at least half of all tumors due to the mutational inactivation of p53.

However, in tumors retaining wild-type p53 the p53-regulated miRNA-encoding genes represent good candidates for being subject to inactivating events. These include loss by deletion or other structural changes as translocations. In addition, downregulation of miRNA expression by epigenetic silencing via CpG methylation and/or deacetylation of promoter regions has been described (reviewed in (Kaur et al. 2016; Gregorova et al. 2021; Morales et al. 2017).

Furthermore, indirect downregulation due to mutations of other upstream regulatory transcription factors and alterations in the miRNA processing machinery has been observed. Another mode of inactivation may be the aberrant expression of a seed-match containing RNA, a so-called competing endogenous RNA (ceRNA), which sequesters the respective miRNA (Salmena et al. 2011). This mechanism was originally observed in plants (Rubio-Somoza et al. 2011). The existence of cancer-relevant ceRNAs in human cells was documented by the identification of RNAs, which regulate expression of the PTEN tumor suppressor via this route (Tay et al. 2011). A further possibility of miRNA inactivation was suggested to occur by

mutation of seed sequences or altered processing of miRNAs. For example, such alterations have been described in lung cancer (Galka-Marciniak et al. 2019). Furthermore, an escape from miRNA action by deletion or mutation of seed-matching sequences is conceivable. Indeed, such alterations have been observed in mRNAs encoding oncogenic factors (Mayr and Bartel 2009; Mayr et al. 2007). For an overview of reported alterations in the p53/miRNA network detected in cancer see Table 7.1.

7.3.1 *Cancer-Specific Alteration of the miR-15/16 Encoding dLEU2 Gene*

The earliest reported genetic inactivation of a miRNA was the observation that the *dLEU2* gene, which is located on chromosome 13q14 and encodes the miR-15a and miR-16-1 miRNAs, is commonly deleted in chronic lymphocytic leukemia (CLL) (Calin et al. 2002). More recently, it was shown that experimental deletion of *miR-15a/16-1* or of the entire *dLEU2* gene predisposes mice to CLL (Klein et al. 2010). Therefore, *dLEU2* is presumably the tumor suppressor gene located in the 13q14 region. Importantly, this study provided the first proof for a *bona fide* tumor suppressor gene function of a miRNA.

7.3.2 *Cancer-Specific Alterations of the* miR-34 *Family*

The *miR-34a* and *miR-34b/c* genes are frequently silenced by CpG methylation in a variety of tumor types (Hermeking 2010; Lodygin et al. 2008; Toyota et al. 2008; Vogt et al. 2011). MiR-34a methylation was initially shown to occur in numerous cell lines derived from different tumor types, including primary prostate cancer and melanoma (Lodygin et al. 2008). Also, the expression of the miR-34 family members miR-34b and miR-34c, which are encoded by a common transcript, is downregulated in many types of cancer. A high frequency of *miR-34b/c* silencing by CpG methylation has been found in colorectal cancer cell lines and colorectal tumor samples (Toyota et al. 2008). We also found CpG methylation of *miR-34b/c* in all 114 cases of primary colorectal cancers analyzed (Vogt et al. 2011). Interestingly, *miR-34b/c* methylation correlated with metastasis and poor survival for several types of cancer (Lujambio et al. 2008). The reintroduction of *miR-34b/c* into cancer cell lines exhibiting *miR-34b/c* silencing inhibited their motility, reduced tumor growth, and inhibited metastasis formation in a xenograft model with an associated downregulation of the respective target genes (e.g., c-MYC, E2F3, and CDK6).

The *miR-34a* gene is located on chromosome 1p36, a region that is commonly deleted in human cancers, as, for example, in neuroblastoma (Thorstensen et al. 2000). Indeed, neuroblastoma often displays loss of *miR-34a* expression (Welch et al. 2007).

Table 7.1 Alterations of p53-regulated miRNAs in human cancer

miRNA-gene	Tumor type	Mechanism	Frequency [%]	*n* =	References
miR-15a / 16-1	Prostate cancer	Deletion	80	20 + 15	Bonci et al. (2008)
	Chronic lymphocytic leukemia/CLL	Germline mutation in the primary precursor	15	75 cancer (+ control: 160 normal)	Calin et al. (2005)
	Chronic lymphocytic leukemia/CLL	Deletion	68, 51	60, 322	Calin et al. (2002); Stilgenbauer et al. (1998)
	Mantle cell lymphoma /MCL	Deletion	55	53	Kohlhammer et al. (2004)
	Mantle cell lymphoma/MCL	Downregulation	71	30	Stilgenbauer et al. (1998)
	Non-small cell lung cancer/NSCLC	Deletion or downregulated	74	23	Bandi et al. (2009); Bandi and Vassella (2011)
	Pituitary tumors [Cushing's disease / CD]	n.d.	–	14 (+ 7controls)	Amaral et al. (2009)
	Ovarian	Downregulation		38	Bhattacharya et al. (2009)
	Non-Hodgkin's lymphoma /NHL	Deletion	43	43	Wada et al. (1999)
	Hodgkin's disease/HD	Deletion	29	7	Wada et al. (1999)
	Multiple myeloma/MM	Downregulation	54	37	Harrison et al. (2003)
	Pituitary adenoma	Deletion	n.d.	20	Bottoni et al. (2005)
	Pancreatic cancer	Downregulation	70	10	Zhang et al. (2010)
	Prostate cancer	Downregulation	100	23	Musumeci et al. (2011)
	Prostate cancer cell lines	Downregulation	–	50	Porkka et al. (2011)
miR-16-2	Progression of prostate carcinogenesis	Downregulation	–	63	Leite et al. (2011)
miR-34a	Non small cell lung cancer/NSCLC	Downregulation	91	23	Bandi and Vassella (2011)
	Acute myeloid leukemia/AML	Hypermethylation	0	20	Chim et al. (2010)
	Non-Hodgkin's lymphoma /NHL	Hypermethylation	18.8	32	Chim et al. (2010)
	Acute lymphoblastic leukemia/ALL	Hypermethylation	0	20	Chim et al. (2010)

	Chronic lymphocytic leukemia/CLL	Hypermethylation	4	50	Chim et al. (2010)
	Chronic myeloid leukemia/CML	Hypermethylation	0	11	Chim et al. (2010)
	Multiple myeloma/MM	Hypermethylation	5,5	55	Chim et al. (2010)
	Colorectal cancer cell lines	Hypermethylation, p53 mutation	23	13	Lodygin et al. (2008)
	Prostate cancer	Hypermethylation	79	24	Lodygin et al. (2008)
	Breast cancer cell lines	Hypermethylation	25	24	Lodygin et al. (2008)
	Kidney cancer cell lines	Hypermethylation	21	14	Lodygin et al. (2008)
	Bladder cancer cell lines	Hypermethylation	33	6	(Lodygin et al. 2008)
	Lung cancer cell lines	Hypermethylation	29	24	Lodygin et al. (2008)
	Melanoma	Hypermethylation	63	32	Lodygin et al. (2008)
miR-34b/c	Gastric cancer	Hypermethylation	70	118	Suzuki et al. (2010)
	Non small cell lung cancer/NSCLC	Hypermethylation		161	Wang et al. (2011)
	Primary melanoma cell lines	Downregulation		2	Migliore et al. (2008)
	Head and neck cancer/H & N	Downregulation	n.d.	10	Cai et al. (2010)
	Colon cancer	Hypermethylation	90	111	Toyota et al. (2008)
	MYC translocation-negative classical Burkitt lymphoma	Downregulation	100	5	Leucci et al. (2008)
miR-34a/b/c	Ovarian cancer	p53 mutation, Hypermethylation	100/72, 62/69	89, 13	Corney et al. (2010); Vogt et al. (2011)
	Malignant pleural mesothelioma/MPM	Hypermethylation	28/85	47	Kubo et al. (2011)
	Colorectal cancer	Hypermethylation	74/99	114	Vogt et al. (2011)
	Pancreatic cancer	Hypermethylation	64/100	11	Vogt et al. (2011)
	Mammary cancer	Hypermethylation	60/90	10	Vogt et al. (2011)
	Urothelial cancer/UC	Hypermethylation	71/57	7	Vogt et al. (2011)
	Renal cell cancer	Hypermethylation	58/100	12	Vogt et al. (2011)
	Soft tissue sarcomas	Hypermethylation	64/ 45	11	Vogt et al. (2011)

(continued)

Table 7.1 (continued)

miRNA-gene	Tumor type	Mechanism	Frequency [%]	*n* =	References
	Esophageal squamous cell carcinoma/ESCC	Hypermethylation	67/41	54	Chen et al. (2012)
miR-107	Head and neck/Oral cancer/HNOC	Downregulation		4	Wong et al. (2008)
	Acute Promyelocytic leukemia/APL	Downregulation		26	Careccia et al. (2009)
	Pancreatic carcinoma cell lines	Hypermethylation		2	Lee et al. (2009)
	Tongue squamous cell carcinoma/TSCC	Downregulation	n.d.	4	Wong et al. (2008)
	Chronic lymphocytic leukemia/CLL	Hypermethylation	n.d.	50	Pallasch et al. (2009)
	Pancreatic cancer	Upregulation		44 (+ 12 controls)	Roldo et al. (2006)
miR-141	Colorectal cancer	Downregulation		10	Baffa et al. (2009)
	Mesenchymal breast cancer Cell lines	Hypermethylation	100	4	Neves et al. (2010)
	Epithelial breast cancer cell lines	Hypermethylation	0	4	Neves et al. (2010)
	Lung cancer	Downregulation		10	Baffa et al. (2009)
	Bladder cancer	Hypermethylation		10 (+5 controls)	Wiklund et al. (2011)
miR-145	Prostate cancer	Hypermethylation, p53 mutation	81	27	Suh et al. (2011)
	Prostate cancer	Downregulation		63	Leite et al. (2011)
miR-192	Colorectal cancer	p53 mutation		34	Karaayvaz et al. (2011)
	Colorectal cancer (MSI)	Downregulation		54 (+ 20 controls)	Earle et al. (2010)
	Multiple myeloma/MM	Hypermethylation		47 (+ 5 controls)	Pichiorri et al. (2010)
miR-194	Colorectal cancer with liver metastasis	p53 mutation, SNP		30	Kahlert et al. (2011)
	Multiple myeloma/MM	Hypermethylation		47 (+ 5 controls)	Pichiorri et al. (2010)
miR-200a	Ovarian cancer	Downregulation		55	Hu et al. (2009)
miR-200b	Colorectal cancer	Loss		30	Kahlert et al. (2011)

	Ovarian cancer	Downregulation		55	Hu et al. (2009)
	Lung cancer	Hypermethylation	25	24 (+ controls)	Tellez et al. (2011)
	Bladder cancer	Hypermethylation		10 (+5 controls)	Wiklund et al. (2011)
miR-200c	Lung cancer	Hypermethylation	29	24 (+ controls)	Tellez et al. (2011)
	Colorectal cancer	Hypermethylation, p53 mutation			Xi et al. (2006)
	Mesenchymal breast cancer cancer cell lines	Hypermethylation	100	4	Neves et al. (2010)
	Epithelial breast cancer cancer cell lines	Hypermethylation	0	4	Neves et al. (2010)
	Lung cancer	Hypermethylation	25, –	24, 69	Ceppi et al. (2010); Tellez et al. (2011)
	Bladder cancer	Hypermethylation		10 (+5 controls)	Wiklund et al. (2011)
miR-215	Multiple myeloma/MM	Hypermethylation		24 (+ controls)	Pichiorri et al. (2010)
	Colorectal cancer	Downregulation		34	Karaayvaz et al. (2011)
miR-429	Colorectal cancer cell lines	Hypermethylation	50	2	Davalos et al. (2011)
	Breast cancer cell lines	Hypermethylation	50	2	Davalos et al. (2011)
	Lung cancer cell lines	Hypermethylation	33	3	Davalos et al. (2011)
	Ovarian cancer	Downregulation		55	Hu et al. (2009)
Exportin 5	Breast cancer	Mutation		441 (+ 479 controls)	Leaderer et al. 2011)
	Hereditary nonpolyposis colon cancer	Downregulation	26	38	Melo et al. 2010)
	Sporadic colon cancer (MSI+)	Downregulation	22	211	Melo et al. (2010)
	Sporadic gastric cancer (MSI+)	Downregulation	28	58	Melo et al. (2010)
	Sporadic endometrial tumors (MSI+)	Downregulation	13	30	Melo et al. (2010)
Dicer/ Drosha	Ovarian cancer	Downregulation	60/51	111	Merritt et al. (2008)
Dicer	Cystic Nephroma, Wilm's tumor	Germline mutation	0	50	Bahubeshi et al. (2011)

(continued)

Table 7.1 (continued)

miRNA-gene	Tumor type	Mechanism	Frequency [%]	*n* =	References
	Pulmonary pediatric cancer	Mutation	91	11 (+ 360 controls)	Hill et al. (2009)
	Lung cancer	Downregulation	67		Karube et al. (2005)
	Colorectal cancer	Upregulation		237	Faber et al. (2011)
	Acute myeloid leukemia/AML	Upregulation	86	71	Martin et al. (2009)

Summary of the reported alterations in p53-regulated miRNAs in cancer. "frequency": relates to the alteration indicated in the third column, "n": number of tumor samples/patients analyzed

7.3.3 *Cancer-Specific Alterations of the* miR-200 *Family*

The miR-200 family encodes a highly conserved group of miRNAs, which controls EMT by downregulating the EMT-inducing transcription factors ZEB1 and ZEB2 (Gregory et al. 2008). The miR-200 family can be subdivided into two clusters: miR-200c and miR-141 (located at chromosome 12p13), and miR-200a, miR-200b, and miR-429 (located at chromosome 1p36). Expression of the miR-200c/141 cluster is frequently silenced by CpG methylation in breast cancer (Neves et al. 2010). Interestingly, a correlation between methylation of the miR-200c promoter and invasiveness was determined in breast cancer cell lines. Downregulation of the miR-200c/141 cluster was also described for breast cancer-initiating cells (Shimono et al. 2009) and Epstein-Barr virus-associated gastric carcinomas (Shinozaki et al. 2010). As mentioned above, loss of 1p36 is a recurrent aberration, especially in neuroblastoma, indicating that there may be two distinct mechanisms that downregulate the expression of the miR-200 family.

7.3.4 *Cancer-Specific Alterations of the* miR-192 *Family*

The p53-regulated miR-192 family is comprised of miR-192, miR-194-2, and miR-215, which induce p21 expression and cell cycle arrest in a p53-dependent manner (Braun et al. 2008). The miR-192 family is downregulated by an unknown mechanism in multiple myeloma (MM), which rarely shows mutation or deletion of p53 (Pichiorri et al. 2010). Reactivation of p53 in MM resulted in re-expression of miR-192, miR-194-2, and miR-215 and downregulation of MDM2, which represents a target of these miRNAs (Pichiorri et al. 2010). Moreover, ectopic expression of miR-192 family members inhibited cell growth, migration, and invasion of MM. Furthermore, the miR-192 family members are downregulated in colon cancer and induce apoptosis and senescence although to a lesser extent than miR-34a (Braun et al. 2008). The mechanism of the miR-192 family downregulation remained unclear in this study, but p53 inactivation (de Krijger et al. 2011) and a single nucleotide polymorphism (SNP) located within the miR-192 precursor (Duan et al. 2007) may contribute to this phenomenon.

7.3.5 *Other p53-Induced miRNAs Inactivated in Cancer*

Recently, the p53-inducible miR-145 was shown to be downregulated by CpG methylation and p53 mutation in prostate cancer samples and cell lines (Suh et al. 2011).

miR-103 and miR-107 were shown to directly target DICER1 mRNA, which encodes a central component of the miRNA processing machinery (Martello et al.

2010). Ectopic expression of miR-103 and miR-107 enhance migration in vitro and allow metastatic dissemination of otherwise nonaggressive cells in vivo, whereas the loss of miR-103/107 opposes migration and metastasis of malignant cells. Moreover, it was shown that high levels of miR-103/107 are associated with metastasis and poor outcome in breast cancer. These findings suggest that the deregulation of the miRNA processing machinery in cancer leads to metastasis and poor outcome, and predicts an anticancer activity of the majority of the miRNAs. In support of this conclusion, DICER1 was characterized as an haploinsufficient tumor suppressor gene in a tumor mouse model (Kumar et al. 2009). Furthermore, decreased expression of DICER1 correlates with poor prognosis in human lung cancer (Karube et al. 2005). Interestingly, the p53 family member p63 transcriptionally controls DICER1 expression. Mutant p53 presumably interferes with this regulation, which leads to a reduction in DICER1 levels and reduces the levels of certain cancer-relevant miRNAs (Su et al. 2010). Moreover, mutant p53 has been shown to inhibit the maturation of a subset of pri-miRNAs via binding and sequestration from the microprocessor complex of the RNA helicases p72/82 (Garibaldi et al. 2016; Suzuki et al. 2009). More recently, p53 was shown to interact with AGO2 and thus regulate the association with AGO2 of a subset of mature miRNAs, such as let-7 family members. Furthermore, specific mutations in p53 decreased the association of let-7 family members with AGO2, thus reducing their activity (Krell et al. 2016).

7.3.6 Mutations in the miRNA Processing Machinery in Cancer

Another possibility of how the abundance of p53-regulated miRNAs could be altered in cancer is to constitutively change the processing of pri-miRNAs to miRNAs by alterations in components of this pathway. For example, mutations of the nuclear export protein Exportin-5 resulted in trapping pre-miRNAs in the nucleus and reduced miRNA processing (Melo et al. 2010). As a result, numerous miRNAs were not fully processed and a diminished inhibition of the respective miRNA targets was detected. Notably, restoration of Exportin-5 functions reversed the impaired export of pre-miRNA and had tumor-suppressive effects. Several studies supported the hypothesis that variations in the expression and mutations of miRNA processing components such as Exportin-5, DROSHA, and DICER1 affect the outcome of breast (Leaderer et al. 2011), ovarian (Merritt et al. 2008), cystic nephroma (Bahubeshi et al. 2011), and pediatric pulmonary cancer (Hill et al. 2009). A large percentage of Wilms tumors, the most common childhood malignancy of the kidney, harbor heterozygous missense mutations in the RNaseIII domain of DROSHA, leading to a global reduction in miRNA levels (Walz et al. 2015; Wegert et al. 2015; Torrezan et al. 2014; Rakheja et al. 2014). In addition, mutations in the dsRNA binding domain of DGCR8, a microprocessor subunit, have also been identified in Wilms tumors (Walz et al. 2015; Wegert et al. 2015; Torrezan et al.

2014). Furthermore, recurrent homozygous deletions of DROSHA have been identified in pineoblastoma (Snuderl et al. 2018).

7.4 Approaches to Study p53-Regulated miRNAs and Their Targets

One currently feasible strategy for a comprehensive, genome-wide identification of p53-regulated miRNAs and their associated target genes is the combination of several approaches depicted in Fig. 7.5. This strategy may in principle also apply to other transcription factors of interest besides p53. These analyses generate a large amount of bioinformatic data, which can be processed with the help of the algorithms indicated in Fig. 7.6. The experimental strategy can be subdivided into two main parts: (1) identification of p53-regulated miRNAs, and (2) identification of target mRNAs of the p53-regulated miRNAs. So far, the studies in this area have rather focused on the identification and characterization of single miRNAs regulated by p53 or they have carried out one type of genome-wide approach, with subsequent confirmation of a limited number of candidates. In the following section, we will describe which approaches have been applied to identify and characterize p53-regulated miRNAs and their associated targets in the past and which lessons have been learned from these analyses.

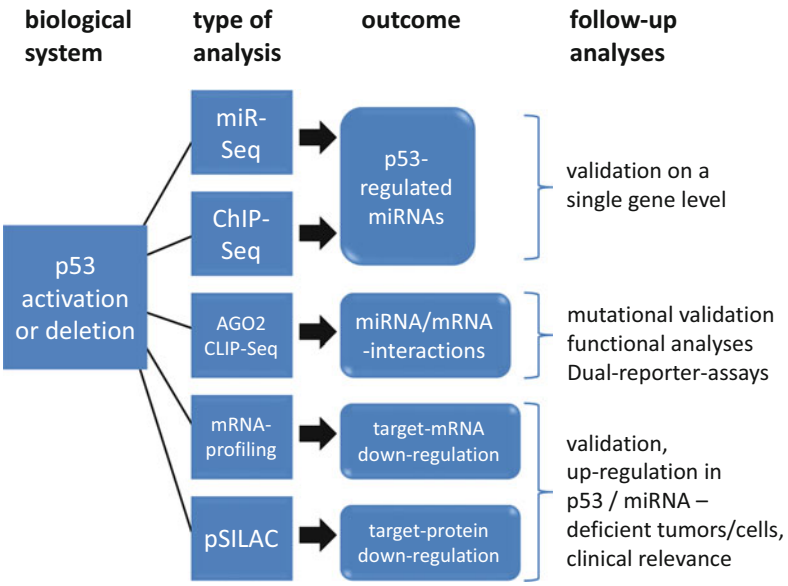


Fig. 7.5 Analysis of p53-regulated miRNAs and their targets. Summary of experimental approaches for the comprehensive identification and characterization of p53-regulated miRNAs. The approaches are described in detail in the text

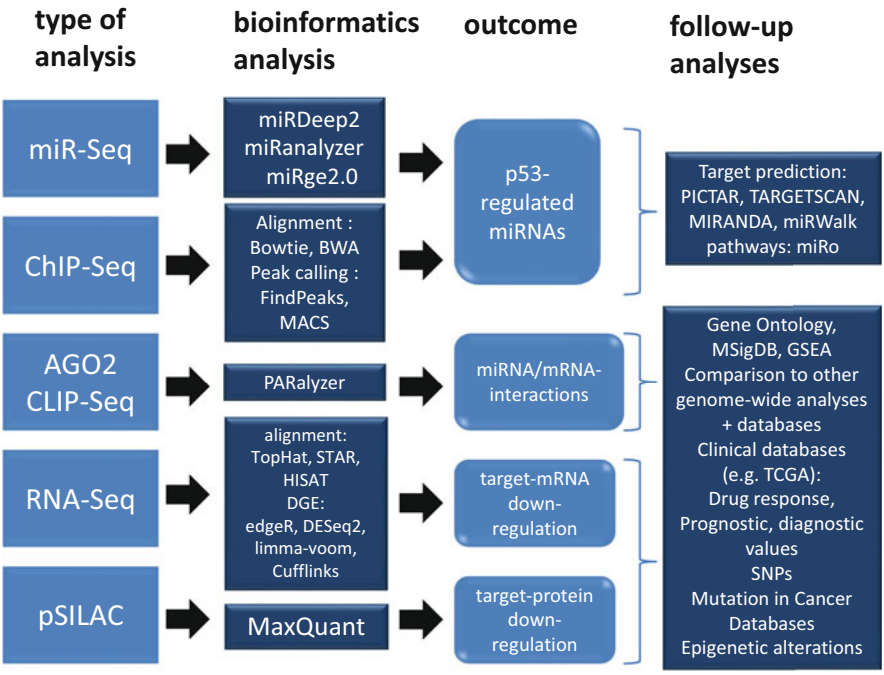


Fig. 7.6 Bioinformatics characterization of p53-regulated miRNAs and their targets. Summary of bioinformatics approaches to facilitate the analyses of data obtained by the experimental analyses described in the main text and in Fig. 7.5 MirDeep2 (Mackowiak 2011), miRanalyzer (Hackenberg et al. 2011) and miRge2.0 (Lu et al. 2018) can be used to analyze miRNA-Seq data. Commonly used mappers for ChIP-Seq data are Bowtie (Langmead et al. 2009) and BWA (Li and Durbin 2009). For peak calling, several tools are available, e.g., Find Peaks (Fejes et al. 2008) or MACS (Feng et al. 2012). AGO2 CLIP-Seq data can be analyzed using PARalyzer (Corcoran et al. 2011). For the mapping of RNA-Seq reads, TopHat2 (Kim et al. 2013), STAR (Dobin et al. 2013), or HISAT (Kim et al. 2015) are widely applied. Several R/Bioconductor-based software packages, such as *EdgeR, DESeq/DESeq2* or *limma-voom,* are commonly used to analyze differential gene expression (DGE) (Law et al. 2014; Anders and Huber 2010; Love et al. 2014; Robinson et al. 2010). Alternatively, mappers such as TopHat can be used in conjunction with CuffLinks to determine differentially expressed genes (Trapnell et al. 2012). For the analysis of large mass spectrometric data sets, MaxQuant (Cox and Mann 2008) can be used. For subsequent pathway analyses, miRo (Lagana et al. 2009), the Molecular Signatures Database (MSigDB) (Liberzon et al. 2015; Liberzon 2014), and gene set enrichment analysis (GSEA) (Subramanian et al. 2005) are commonly used tools

7.4.1 Identification of p53-Regulated miRNAs

In order to experimentally identify p53-regulated miRNAs, cellular systems in which p53 activity can be turned on using conditional systems or pharmacological p53 activators should be employed. Endogenous p53 can either be activated by the addition of DNA damaging substances or by specific p53-activators such as Nutlin-

3a. Isogenic cells with and without wild-type p53 should be treated in parallel in order to identify p53-dependent regulations. For example, the colon cancer cell lines HCT 116 exhibiting either wild-type p53 expression or p53 knockout are useful for this purpose (Bunz et al. 1998). Alternatively, the miRNA expression in tissues of p53 knockout mice or derived cells, e.g., mouse embryonic fibroblasts (MEFs), represent useful systems to identify p53-mediated miRNA regulations, as documented previously (He et al. 2007a).

A more specific activation of p53 can be achieved using ectopic expression of p53. However, certain posttranslational modifications of p53 induced by treatment with DNA damaging agents or pharmacological activators do not occur hereby and therefore, differences in the pattern of miRNAs regulated by p53 may occur when compared to activation of p53 by stressors such as oncogene activation and DNA damaging agents. In the past, we have used an episomal, doxycyclin-inducible expression system to re-express p53 in p53-deficient H1299 lung cancer cells (Tarasov et al. 2007). Since p53 may enhance the synthesis of miRNAs by directly influencing pre-miRNA processing, the detection of differential expression of the mature miRNA is not sufficient to deduce a direct transcriptional regulation of the corresponding pri-miRNA by p53 (Suzuki et al. 2009). Therefore, it is advantageous to obtain both miRNA and pri-miRNA profiles simultaneously.

Differential expression of mature miRNAs upon p53 activation can be monitored using specifically designed miRNA microarrays. A number of commercially available microarray platforms can be used for this purpose, such as the Human miRNA Microarray Release 21.0 (Agilent) or the Genechip miRNA 4.0 array (Affymetrix).

Several previous studies have used microarrays to identify miR-34 and miR-215/miR-192 as direct p53 targets. A miRdicator array was used to identify *miR-34*a as a p53 target gene (Raver-Shapira et al. 2007), a 4X2K microarray (CombiMatrix) that contained probes against mouse miRNAs identified *miR-34b/c* as a p53 target gene (Corney et al. 2007) and customized miRNA arrays were used to detect *miR-34a* (Chang et al. 2007) and *miR-192/miR-215* (Braun et al. 2008) as p53 target genes. For example, two studies employed miRNA microarrays to identify members of the miR-200 family as p53 targets (Chang et al. 2011; Kim et al. 2011c).

For more focused approaches, pre-designed PCR panels covering known disease-relevant miRNAs such as the miRCURY LNA miRNA Focus PCR Panels (QIAGEN) or TaqMan Arrays (ThermorFisher) can also be used. In these assays, the induction of mature miRNAs after p53 activation can be measured by stem-loop RT-qPCR. Hannon and colleagues used a panel of 145 TaqMan assays to monitor changes in mature miRNA levels after p53 activation (He et al. 2007a). This approach may also be used to verify the microarray expression data at the level of individual, processed miRNAs. In order to determine whether p53 regulates miRNA expression at the transcriptional level, induction of the pri-miRNA transcript can be measured using total mRNA preparations after reverse transcription into cDNAs and standard real-time quantitative PCR (qPCR).

A subset of miRNAs lies within intronic sequences of host genes, and therefore differential expression of the host mRNAs can in principle be monitored by standard gene expression arrays used for mRNAs. However, induction of the primary host

transcript does not necessarily lead to a significant induction of the mature miRNA. Therefore, the induction of the mature miRNA should be validated by stem-loop RT-qPCR assays. The above-mentioned methods have in common that they only detect previously known miRNAs.

For the unbiased detection of all miRNA expressed in a certain state, several Next Generation Sequencing (NGS) based approaches are currently being used. Small RNAs are isolated, ligated to adapters, reverse transcribed, and amplified to generate libraries, which may be analyzed using different NGS platforms, e.g., HiSeq or MiSeq (Illumina), IonTorrent (ThermoFisher), or the SOLID system (Applied Biosystems) (reviewed in (Liu 2018; Hu et al. 2017). The adapters often contain distinct bar codes, which allow multiplexing of several samples in one sequencing run generating up to several hundred million reads. The coverage which can be achieved by these analyses is presumably close to complete. In 2007, we applied a 454-sequencing approach to identify *miR-34a* as direct p53 target (Tarasov et al. 2007). At that time only ~200.000 sequencing reads per run were reached, but were sufficient to identify many of the miRNAs with the most pronounced regulation by p53.

7.4.2 Confirmation of Direct Regulation by p53 Using ChIP Approaches

The detection of p53 occupancy at the respective promoters of the genes encoding p53-regulated pri-miRNAs or other precursor mRNAs can be achieved by chromatin immunoprecipitation (ChIP)-based techniques. These can either be performed on a gene-by-gene basis using qPCR-ChIP or on a genome-wide level by coupling ChIP with techniques such as NGS, SAGE, or hybridization to a promoter array. The disadvantage of the latter is a bias toward previously characterized promoters that are spotted.

The consensus sequence necessary for p53 binding consists of two copies of the RRRCWWGYYY motif separated by a small spacer of 0-21 nucleotides (R = pyrimidine; Y = purine; W = A/T; see also Fig. 7.1). However, among the validated p53 response elements identified in p53 target gene promoters, the majority display slight deviations from the consensus sequence, indicating a degeneracy in p53's binding requirements. Based on the consensus motif, potential p53 binding sites can be predicted using a variety of search algorithms. For example, in the context of p53-regulated miRNAs, the p53MH algorithm (Hoh et al. 2002) and the MatInspector software (Genomatix) have been applied to identify p53 binding sites in the promoters of miRNA-encoding genes. The P53MH algorithm was used to identify a p53 binding site in the *miR-34b/c* promoter (Corney et al. 2007) and in the *miR-194-1/miR-215* cluster (Braun et al. 2008), whereas the two p53 binding sites in the *miR-145* promoter were identified using the MatInspector software (Sachdeva et al. 2009).

Initially, binding of p53 to the identified binding site was experimentally tested in vitro by gel shift assays. Furthermore, in order to test the requirement of the p53 response element, a genomic region encompassing the p53 binding site or its mutant version can be placed upstream of a *luciferase* ORF or an equivalent reporter gene. The responsiveness of these constructs to p53 is interrogated by co-transfection with p53-encoding plasmids into mammalian cells and a subsequent reporter assay. In order to test whether p53 binds to the identified binding site in a native chromatin environment in vivo, chromatin immunoprecipitation assays have to be performed. This can either be done on a single gene basis by ChIP followed by semiquantitative PCR or qPCR or p53 binding sites can also be identified on a genome-wide scale. In the initial genome-wide binding studies, immunoprecipitated DNA from the ChIP experiment was hybridized into high-density oligonucleotide tiling arrays (ChIP-on-Chip). Cawley et al. used a ChIP-on-Chip approach to map p53 binding sites on human chromosomes 21 and 22 and identified 48 high-confidence sites (Cawley et al. 2004). These results suggested the existence of ~1600 putative p53 sites in the human genome. When the same approach was applied to the complete genome, 1546 p53 binding sites were identified in actinomycin D-treated U2OS cells (Smeenk et al. 2008).

The ChIP-PET method is an extension of the ChIP-on-Chip approach and is related to SAGE (Velculescu et al. 1995). Short tags derived from immunoprecipitated DNA fragments are converted into a DNA library. After further ligations, the paired-end ditags form concatemeres, which are subjected to capillary sequencing. The obtained tag-sequences reads are subsequently mapped to the genome and quantified. Wei et al. used the ChIP-PET method to monitor p53 binding across the whole genome and identified more than 500 high-confidence p53 binding sites (Wei et al. 2006). This resource was used by other laboratories to identify p53 binding sites in the *miR-34a* and *miR-34b/c* promoters (Bommer et al. 2007; Raver-Shapira et al. 2007).

The methods mentioned above have been largely replaced by a combination of ChIP and NGS (ChIP-Seq). Since the latest generation of sequencing devices achieve up to several billion reads in one run, it is possible to multiplex several libraries representing, for example, different time points and experimental replicas in one single sequencing run. The identification of occupied p53 binding sites in the genome may be combined with detection of histone modifications indicating active transcription units and enhancers. This allows the assignment of orphan miRNAs to active promoters present in their vicinity. Furthermore, the results obtained using the expression studies described above have to be compared to the DNA binding patterns of p53 in a genome-wide manner using bioinformatic approaches.

7.4.3 Identification of miRNA Targets

Having obtained a set of p53-regulated miRNAs, the next step is to identify the physiologically relevant target mRNAs of these miRNAs. We suggest the systematic

identification of miRNA-regulated target genes following p53 induction by an integrated approach that involves:

A. Identification and mapping of miRNA binding sites using biochemical techniques involving RISC isolation.
B. Testing the functionality of these binding sites in the regulation of their respective target mRNAs using either microarrays or NGS as well as dual-reporter assays.
C. Proteomic approaches to measure changes in target abundance on the protein level indicate translational regulation in cases without a decrease in the corresponding mRNA.

Similar to the identification of p53-induced miRNAs in part 1, these approaches should be ideally performed in parallel as they complement each other. The identification and mapping of miRNA binding sites in the transcriptome provide information as to whether a miRNA directly binds to its cognate target mRNA, but does not provide information about the regulation of the bound mRNA. Conversely, microarray and proteomic approaches provide information on the regulation of a given mRNA or protein, but do not per se distinguish between direct and indirect targets. Therefore, a combined approach that maps binding sites of p53-regulated miRNAs in the transcriptome and validates the functionality of these binding sites regarding target regulation may comprehensively uncover the network of protein expression that is regulated by p53-induced miRNAs.

MiRNAs typically regulate their targets via association of a ~7 nucleotide stretch, the so-called seed-sequence, located in their 5′-portion with a complementary sequence in the 3´-UTR of the target mRNA. Additional base pairing may occur via nucleotides in the middle and 3′-portion of the miRNA. Since miRNAs usually pair imperfectly with their respective target mRNAs, the number of theoretically possible targets is typically large and presumably most of the predicted targets are not significantly regulated by the respective miRNA. Several bioinformatic algorithms have been developed to predict miRNA targets with the intention to reduce the rate of false-positive predictions by incorporating features such as conservation between species. However, even these algorithms often predict hundreds of target mRNAs for a particular miRNA, of which not all are necessarily physiologically relevant targets.

Numerous miRNA target prediction tools have been developed in recent years (reviewed in Roberts and Borchert (2017), Loganantharaj and Randall (2017)). Here, we focus on the most commonly used algorithms and their underlying principles of prediction. Due to differences in the parameters used to weigh individual features involved in miRNA/mRNA interaction, different target prediction algorithms often result in only partially overlapping sets of predicted target genes. Algorithms like TargetScan and Pictar (Friedman et al. 2009; Krek et al. 2005) place more weight on perfect, evolutionarily conserved seed matches, whereas PITA, RNA22, and RNAhybrid (Kertesz et al. 2007; Miranda et al. 2006; Rehmsmeier et al. 2004) prioritize the ΔG of the miRNA/mRNA duplex and the accessibility of the site within the mRNA. Although algorithms like Target Scan and Pictar have been

shown to have high predictive power when tested on experimentally obtained proteomic data (Alexiou et al. 2009; Baek et al. 2008; Selbach et al. 2008), they may be less useful in the prediction of miRNA target sites that lack a perfect seed-sequence, are not evolutionarily conserved, or lie outside the 3´-UTR of the target gene. Therefore, the combined use of several different algorithms may be helpful to identify target mRNAs of a given miRNA. For example, a useful resource combining predictions of several algorithms is the miRWalk database (Dweep et al. 2014; Sticht et al. 2018). In addition, ComiR (Combinatorial miRNA targeting) applies a combination of several prediction algorithms to predict mRNA coordinately regulated by several miRNAs, which helps to improve miRNA target prediction (Coronnello and Benos 2013; Bertolazzi et al. 2020).

The sets of predicted target mRNAs generated by different algorithms are typically used to filter sets of differentially regulated genes that were identified by experimental perturbation of miRNA function. This is followed by unbiased genome- or proteome-wide measurements of changes in mRNA or protein abundance. As outlined in Fig. 7.5, miRNA binding sites can be mapped by isolation of RISC-bound miRNA target mRNAs. This is typically accomplished by immunoprecipitation of RISC components such as Ago2, which can either be done via endogenous proteins or ectopically expressed epitope-tagged versions of the respective proteins (Beitzinger et al. 2007; Hendrickson et al. 2008; Karginov et al. 2007). The RISC/mRNA/miRNA complexes are precipitated and the associated mRNAs are identified either by hybridization to microarrays or by NGS technologies. However, this method does not directly lead to the identification of the actual miRNA binding site, since the entire RISC-bound mRNA is immunoprecipitated and sequenced.

An improved version of these initial approaches is high-throughput sequencing of RNAs isolated by cross-linking and immunoprecipitation (HITS-CLIP (Chi et al. 2009): miRNA-bound RNAs are cross-linked to RISC by UV irradiation. The RISC/miRNA/mRNA complex is then immunoprecipitated with antibodies against RISC components such as Ago2. A RNAse-digest eliminates all mRNA fragments not protected by the RISC/miRNA complex. In another version of an AGO2-IP-based approach named photoactivatable-ribonucleoside-enhanced cross-linking and immunoprecipitation (PAR-CLIP) (Hafner et al. 2010), cells are cultured with the photo-reactive 4-thiouridine before UV-cross-linking. 4-thiouridine is incorporated into the cellular RNA during transcription and leads to improved cross-linking efficiencies. Since 4-thiouridine results in C-to-T transitions during reverse transcription, it helps to map the exact miRISC-bound position on the mRNA.

All occupied mRNA sites, presumably miRNA matching regions, are then determined by NGS. Therefore, these approaches have to be combined with specific bioinformatics workflows to extract the associated sequence motifs from the sequencing data (Chakrabarti et al. 2018). Following this approach, information is obtained not only regarding the bound mRNA target but also the miRNA matching sequence, which allows to deduce the putative identity of the miRNA. In the case of p53-induced miRNAs, this miRNA should be among those which are detected at increased levels after p53 activation. A PAR-CLIP approach has been applied to

identify the mRNA targets of DNA damage-induced miRNAs in HCT116 *TP53*$^{+/+}$ and *TP53*$^{-/-}$ cells (Krell et al. 2016). This study identified a role for p53 in the association of let-7 miRNA family members with AGO2 and its target mRNAs.

Since all these approaches essentially rely on the isolation of the RISC complex, all miRNAs and their bound mRNA targets associated with RISC will be identified, which means that identification of mRNA targets of a particular miRNA from the obtained sequence data largely depends on the subsequent extraction of sequence features associated with the miRNA. This includes the presence of a hexameric seed sequence or the presence of other sequence features predicted to be targeted by miRNAs by algorithms such as PITA or RNA22. An alternative involves the use of biotinylated miRNAs that can be purified together with RISC in a tandem affinity purification approach (Orom and Lund 2007, 2010).

As explained above, information on the miRNA binding site does not automatically mean that this particular binding site is physiologically relevant for target regulation. Therefore, changes in either mRNA or protein abundance have to be determined following perturbation of miRNA expression. Experimental studies to identify target mRNAs of p53-regulated miRNAs often involve gain-of-function approaches using ectopic expression of miRNAs either by transfection of synthetic pre-miRNA molecules or inducible expression of pri-miRNA transcripts (Bommer et al. 2007; Kaller et al. 2011; Tarasov et al. 2007). In the opposite loss-of-function approach, synthetic miRNA inhibitors (antagomiRs) can be used to block miRNA function. Alternatively, and more elegantly, knockout cell lines for individual miRNAs can be used to address this question. In recent years, the progress in CRISPR/Cas9-mediated genome engineering has greatly facilitated the generation of such miRNA knockout cell lines (Chang et al. 2016; Lataniotis et al. 2017). In addition, the HCT116 DICER^{ex5}, a human colorectal cancer cell line with an engineered hypomorphic DICER allele (Cummins et al. 2006), has been used to characterize targets of p53-regulated miRNAs (He et al. 2007a, b; Georges et al. 2008).

A number of studies applied microarrays to identify targets of p53-induced miRNAs, such as miR-34 (Bommer et al. 2007; Chang et al. 2007; He et al. 2007a; Tazawa et al. 2007) and mir-216/mir-192 (Georges et al. 2008). Microarray approaches are limited as they cannot detect miRNA targets that are solely regulated at the level of translational repression. On the other hand, assuming that miRNAs in most cases only cause modest decreases in protein translation (Selbach et al. 2008; Baek et al. 2008), the miRNA-mediated regulation of proteins with long half-lives may not be detected by measuring steady-state protein levels using standard proteomic quantification as SILAC (***s***table ***i***sotope ***l***abeling by ***a***mino acids in ***c***ell culture) (Ong et al. 2002). This problem was solved by the introduction of pSILAC (***p***ulsed SILAC), which facilitated the quantification of differences in protein translation rates caused by miRNAs (Selbach et al. 2008). In this approach, induction of miRNA expression is followed by a pulse of isotope-labeled amino acids which are incorporated into newly synthesized proteins. Subsequent mass spectrometric analysis of the proteome, therefore, allows to detect of changes in protein translation rates after miRNA expression. We have applied this approach to identify miRNA

target genes after ectopic expression of miR-34a or after induction of p53. Notably, numerous of the identified miR-34a targets were confirmed in a miRNA capture approach using biotinylated miR-34a as a bait (Lal et al. 2011). Other quantitative proteomic methods like isotope-coded affinity tag (ICAT)-labeling following transfection with miR-34a have been used to identify miRNA targets (Chen et al. 2011). One major drawback of all proteomic methods is their still limited ability to cover the entire proteome of the cell, as well as their strong bias for highly expressed proteins.

All genome- or proteome-wide methods to identify miRNA targets require further validation such as qPCR or Western blot analysis to verify that a given mRNA or protein is indeed regulated following miRNA induction.

Direct regulation by a miRNA is determined in dual-reporter assays. For this, the 3´-UTR of the putative target mRNA is placed downstream of a firefly luciferase reporter gene. This reporter-construct is co-transfected either with miRNA mimics or miRNA inhibitors and a Renilla luciferase vector for standardization. In case of specific, direct regulation the 3´-UTR reporter is repressed by ~20-80%. In order to map and validate the seed-matching sequences these should be mutated in the context of its 3´-UTR sequence. The resulting constructs should ideally show resistance toward the respective miRNA.

7.4.4 Follow-Up Analysis

Once p53-regulation of miRNAs and their targets have been confirmed, numerous additional analyses are possible to interrogate the physiological and pathophysiological relevance of the identified regulations. In recent years, CRISPR/Cas9-mediated genome engineering has made it possible to rapidly generate miRNA knockout cell lines or cell lines with specific alterations in individual miRNA binding sites. Furthermore, the relevance of the respective miRNAs can be tested in knockout mice in combination with tumor mouse models. For this purpose, a collection of ES cell lines with deletion of 392 miRNAs was generated to facilitate the rapid generation of knockout mice (Prosser et al. 2011). Furthermore, the inactivation of the respective miRNA encoding genes by CpG methylation or mutations in different types of cancer may be analyzed and correlated with the putative upregulation of miRNA targets. In the recent years, publicly available datasets of cancer patient cohorts, e.g., from The Cancer Genome Atlas (TCGA), combining miRNA and mRNA expression, have been increasingly used to interrogate the role of p53 mutations on miRNA and miRNA target expression and to determine correlations with mutations, epigenetic changes and clinical data (Donehower et al. 2019).

7.4.5 Outlook

In the future technological developments may result in an increase in sensitivities of mass spectrometry analyses which may facilitate similar coverage rates of proteomic quantifications as are now reached by RNA sequencing-based approaches. Furthermore, the integration of different bioinformatic platforms into a common program for mRNA/miRNA/DNA binding and protein quantification will make integrated analyses less complicated and laborious. For example, BioVLAB-MMIA-NGS has been introduced as an integrated analysis system for both miRNA and mRNA sequencing data to identify relevant miRNA-mediated mRNA regulations (Chae et al. 2015). Another useful tool would be a comprehensive ontology-like database for miRNA functions and targets. The miRo website is an example of such a tool (Lagana et al. 2009).

High-throughput single cell (sc) expression analyses, such as scRNA-Seq and sc-proteomics, will allow to determine the regulations of mRNA and protein expression, e.g., by p53-induced miRNAs on the cellular level. This will facilitate the evaluation of cell type-specific expression as well as heterogeneity of expression in normal tissue and tumor (Vistain and Tay 2021; Wang et al. 2019; Garg and Sharp 2016; Schmiedel et al. 2015; Kumar et al. 2014; Marx 2019). Taken together, these possibilities will hopefully lead to the translation of knowledge about the p53/miRNA network into diagnostic and therapeutic applications.

Acknowledgments We thank the members of the Hermeking Lab for discussions. Work in HH's lab is supported by the Wilhelm-Sander-Stiftung, the Rudolf-Bartling-Stiftung, the Deutsche Krebshilfe, and the Deutsche Forschungsgemeinschaft (DFG).

References

Afanasyeva EA, Mestdagh P, Kumps C, Vandesompele J, Ehemann V, Theissen J, Fischer M, Zapatka M, Brors B, Savelyeva L, Sagulenko V, Speleman F, Schwab M, Westermann F (2011) MicroRNA miR-885-5p targets CDK2 and MCM5, activates p53 and inhibits proliferation and survival. Cell Death Differ 18(6):974–984

Alexiou P, Maragkakis M, Papadopoulos GL, Reczko M, Hatzigeorgiou AG (2009) Lost in translation: an assessment and perspective for computational microRNA target identification. Bioinformatics 25(23):3049–3055

Amaral FC, Torres N, Saggioro F, Neder L, Machado HR, Silva WA Jr, Moreira AC, Castro M (2009) MicroRNAs differentially expressed in ACTH-secreting pituitary tumors. J Clin Endocrinol Metab 94(1):320–323

Anders S, Huber W (2010) Differential expression analysis for sequence count data. Genome Biol 11(10):R106

Baek D, Villen J, Shin C, Camargo FD, Gygi SP, Bartel DP (2008) The impact of microRNAs on protein output. Nature 455(7209):64–71

Baffa R, Fassan M, Volinia S, O'Hara B, Liu CG, Palazzo JP, Gardiman M, Rugge M, Gomella LG, Croce CM, Rosenberg A (2009) MicroRNA expression profiling of human metastatic cancers identifies cancer gene targets. J Pathol 219(2):214–221

Bahubeshi A, Tischkowitz M, Foulkes WD (2011) miRNA processing and human cancer: DICER1 cuts the mustard. Sci Transl Med 3(111):111ps146

Bandi N, Vassella E (2011) miR-34a and miR-15a/16 are co-regulated in non-small cell lung cancer and control cell cycle progression in a synergistic and Rb-dependent manner. Mol Cancer 10:55

Bandi N, Zbinden S, Gugger M, Arnold M, Kocher V, Hasan L, Kappeler A, Brunner T, Vassella E (2009) miR-15a and miR-16 are implicated in cell cycle regulation in a Rb-dependent manner and are frequently deleted or down-regulated in non-small cell lung cancer. Cancer Res 69(13): 5553–5559

Beitzinger M, Peters L, Zhu JY, Kremmer E, Meister G (2007) Identification of human microRNA targets from isolated argonaute protein complexes. RNA Biol 4(2):76–84

Bertolazzi G, Benos PV, Tumminello M, Coronnello C (2020) An improvement of ComiR algorithm for microRNA target prediction by exploiting coding region sequences of mRNAs. BMC Bioinformatics 21(Suppl 8):201

Bhattacharya R, Nicoloso M, Arvizo R, Wang E, Cortez A, Rossi S, Calin GA, Mukherjee P (2009) MiR-15a and MiR-16 control Bmi-1 expression in ovarian cancer. Cancer Res 69(23): 9090–9095

Bohlig L, Friedrich M, Engeland K (2011) p53 activates the PANK1/miRNA-107 gene leading to downregulation of CDK6 and p130 cell cycle proteins. Nucleic Acids Res 39(2):440–453

Bommer GT, Gerin I, Feng Y, Kaczorowski AJ, Kuick R, Love RE, Zhai Y, Giordano TJ, Qin ZS, Moore BB, MacDougald OA, Cho KR, Fearon ER (2007) p53-mediated activation of miRNA34 candidate tumor-suppressor genes. Curr Biol 17(15):1298–1307

Bonci D, Coppola V, Musumeci M, Addario A, Giuffrida R, Memeo L, D'Urso L, Pagliuca A, Biffoni M, Labbaye C, Bartucci M, Muto G, Peschle C, De Maria R (2008) The miR-15a-miR-16-1 cluster controls prostate cancer by targeting multiple oncogenic activities. Nat Med 14(11): 1271–1277

Bottoni A, Piccin D, Tagliati F, Luchin A, Zatelli MC, Degli Uberti EC (2005) miR-15a and miR-16-1 down-regulation in pituitary adenomas. J Cell Physiol 204(1):280–285

Bou Kheir T, Futoma-Kazmierczak E, Jacobsen A, Krogh A, Bardram L, Hother C, Gronbaek K, Federspiel B, Lund AH, Friis-Hansen L (2011) miR-449 inhibits cell proliferation and is down-regulated in gastric cancer. Mol Cancer 10:29

Brabletz T (2012) MiR-34 and SNAIL: another double-negative feedback loop controlling cellular plasticity/EMT governed by p53. Cell Cycle 11(2):215

Braun CJ, Zhang X, Savelyeva I, Wolff S, Moll UM, Schepeler T, Orntoft TF, Andersen CL, Dobbelstein M (2008) p53-responsive micrornas 192 and 215 are capable of inducing cell cycle arrest. Cancer Res 68(24):10094–10104

Bunz F, Dutriaux A, Lengauer C, Waldman T, Zhou S, Brown JP, Sedivy JM, Kinzler KW, Vogelstein B (1998) Requirement for p53 and p21 to sustain G2 arrest after DNA damage. Science 282(5393):1497–1501

Cai KM, Bao XL, Kong XH, Jinag W, Mao MR, Chu JS, Huang YJ, Zhao XJ (2010) Hsa-miR-34c suppresses growth and invasion of human laryngeal carcinoma cells via targeting c-met. Int J Mol Med 25(4):565–571

Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, Aldler H, Rattan S, Keating M, Rai K, Rassenti L, Kipps T, Negrini M, Bullrich F, Croce CM (2002) Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A 99(24):15524–15529

Calin GA, Ferracin M, Cimmino A, Di Leva G, Shimizu M, Wojcik SE, Iorio MV, Visone R, Sever NI, Fabbri M, Iuliano R, Palumbo T, Pichiorri F, Roldo C, Garzon R, Sevignani C, Rassenti L, Alder H, Volinia S, Liu CG, Kipps TJ, Negrini M, Croce CM (2005) A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N Engl J Med 353(17):1793–1801

Careccia S, Mainardi S, Pelosi A, Gurtner A, Diverio D, Riccioni R, Testa U, Pelosi E, Piaggio G, Sacchi A, Lavorgna S, Lo-Coco F, Blandino G, Levrero M, Rizzo MG (2009) A restricted signature of miRNAs distinguishes APL blasts from normal promyelocytes. Oncogene 28(45): 4034–4040

Cawley S, Bekiranov S, Ng HH, Kapranov P, Sekinger EA, Kampa D, Piccolboni A, Sementchenko V, Cheng J, Williams AJ, Wheeler R, Wong B, Drenkow J, Yamanaka M, Patel S, Brubaker S, Tammana H, Helt G, Struhl K, Gingeras TR (2004) Unbiased mapping of transcription factor binding sites along human chromosomes 21 and 22 points to widespread regulation of noncoding RNAs. Cell 116(4):499–509

Ceppi P, Mudduluru G, Kumarswamy R, Rapa I, Scagliotti GV, Papotti M, Allgayer H (2010) Loss of miR-200c expression induces an aggressive, invasive, and chemoresistant phenotype in non-small cell lung cancer. Mol Cancer Res 8(9):1207–1216

Chae H, Rhee S, Nephew KP, Kim S (2015) BioVLAB-MMIA-NGS: microRNA-mRNA integrated analysis using high-throughput sequencing data. Bioinformatics 31(2):265–267

Chakrabarti AM, Haberman N, Praznik A, Luscombe NM, Ule J (2018) Data science issues in studying protein–RNA interactions with CLIP technologies. Annual Review of Biomedical Data Science 1(1):235–261

Chang CJ, Chao CH, Xia W, Yang JY, Xiong Y, Li CW, Yu WH, Rehman SK, Hsu JL, Lee HH, Liu M, Chen CT, Yu D, Hung MC (2011) p53 regulates epithelial-mesenchymal transition and stem cell properties through modulating miRNAs. Nat Cell Biol 13(3):317–323

Chang H, Yi B, Ma R, Zhang X, Zhao H, Xi Y (2016) CRISPR/cas9, a novel genomic tool to knock down microRNA in vitro and in vivo. Sci Rep 6:22312

Chang TC, Wentzel EA, Kent OA, Ramachandran K, Mullendore M, Lee KH, Feldmann G, Yamakuchi M, Ferlito M, Lowenstein CJ, Arking DE, Beer MA, Maitra A, Mendell JT (2007) Transactivation of miR-34a by p53 broadly influences gene expression and promotes apoptosis. Mol Cell 26(5):745–752

Chen QR, Yu LR, Tsang P, Wei JS, Song YK, Cheuk A, Chung JY, Hewitt SM, Veenstra TD, Khan J (2011) Systematic proteome analysis identifies transcription factor YY1 as a direct target of miR-34a. J Proteome Res 10(2):479–487

Chen X, Hu H, Guan X, Xiong G, Wang Y, Wang K, Li J, Xu X, Yang K, Bai Y (2012) CpG island methylation status of miRNAs in esophageal squamous cell carcinoma. Int J Cancer 130(7): 1607–1613

Cheok CF, Verma CS, Baselga J, Lane DP (2011) Translating p53 into the clinic. Nat Rev Clin Oncol 8(1):25–37

Chi SW, Zang JB, Mele A, Darnell RB (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature 460(7254):479–486

Chim CS, Wong KY, Qi Y, Loong F, Lam WL, Wong LG, Jin DY, Costello JF, Liang R (2010) Epigenetic inactivation of the miR-34a in hematological malignancies. Carcinogenesis 31(4): 745–750

Choi YJ, Lin CP, Ho JJ, He X, Okada N, Bu P, Zhong Y, Kim SY, Bennett MJ, Chen C, Ozturk A, Hicks GG, Hannon GJ, He L (2011) miR-34 miRNAs provide a barrier for somatic cell reprogramming. Nat Cell Biol 13(11):1353–1360

Chou J, Kaller M, Jaeckel S, Rokavec M, Hermeking H (2022) AP4 suppresses DNA damage, chromosomal instability and senescence via inducing MDC1/Mediator of DNA damage checkpoint 1 and repressing MIR22HG/miR-22-3p. Mol Cancer 21(1):120

Corcoran DL, Georgiev S, Mukherjee N, Gottwein E, Skalsky RL, Keene JD, Ohler U (2011) PARalyzer: definition of RNA binding sites from PAR-CLIP short-read sequence data. Genome Biol 12(8):R79

Corney DC, Flesken-Nikitin A, Godwin AK, Wang W, Nikitin AY (2007) MicroRNA-34b and MicroRNA-34c are targets of p53 and cooperate in control of cell proliferation and adhesion-independent growth. Cancer Res 67(18):8433–8438

Corney DC, Hwang CI, Matoso A, Vogt M, Flesken-Nikitin A, Godwin AK, Kamat AA, Sood AK, Ellenson LH, Hermeking H, Nikitin AY (2010) Frequent downregulation of miR-34 family in human ovarian cancers. Clin Cancer Res 16(4):1119–1128

Coronnello C, Benos PV (2013) ComiR: combinatorial microRNA target prediction tool. Nucleic Acids Res 41(Web Server issue):W159–W164

Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26(12): 1367–1372

Cummins JM, He Y, Leary RJ, Pagliarini R, Diaz LA Jr, Sjoblom T, Barad O, Bentwich Z, Szafranska AE, Labourier E, Raymond CK, Roberts BS, Juhl H, Kinzler KW, Vogelstein B, Velculescu VE (2006) The colorectal microRNAome. Proc Natl Acad Sci U S A 103(10): 3687–3692

Davalos V, Moutinho C, Villanueva A, Boque R, Silva P, Carneiro F, Esteller M (2011) Dynamic epigenetic regulation of the microRNA-200 family mediates epithelial and mesenchymal transitions in human tumorigenesis. Oncogene 31(16):2062–2074

de Krijger I, Mekenkamp LJ, Punt CJ, Nagtegaal ID (2011) MicroRNAs in colorectal cancer metastasis. J Pathol 224(4):438–447

Derheimer FA, Kastan MB (2010) Multiple roles of ATM in monitoring and maintaining DNA integrity. FEBS Lett 584(17):3675–3681

di Gennaro A, Damiano V, Brisotto G, Armellin M, Perin T, Zucchetto A, Guardascione M, Spaink HP, Doglioni C, Snaar-Jagalska BE, Santarosa M, Maestro R (2018) A p53/miR-30a/ZEB2 axis controls triple negative breast cancer aggressiveness. Cell Death Differ 25(12):2165–2180

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29(1):15–21

Donehower LA, Soussi T, Korkut A, Liu Y, Schultz A, Cardenas M, Li X, Babur O, Hsu TK, Lichtarge O, Weinstein JN, Akbani R, Wheeler DA (2019) Integrated analysis of TP53 gene and Pathway alterations in the cancer genome atlas. Cell Rep 28(5):1370–1384 e1375

Duan R, Pak C, Jin P (2007) Single nucleotide polymorphism associated with mature miR-125a alters the processing of pri-miRNA. Hum Mol Genet 16(9):1124–1131

Dweep H, Gretz N, Sticht C (2014) miRWalk database for miRNA-target interactions. Methods Mol Biol 1182:289–305

Earle JS, Luthra R, Romans A, Abraham R, Ensor J, Yao H, Hamilton SR (2010) Association of microRNA expression with microsatellite instability status in colorectal adenocarcinoma. J Mol Diagn 12(4):433–440

Fabbri M, Bottoni A, Shimizu M, Spizzo R, Nicoloso MS, Rossi S, Barbarotto E, Cimmino A, Adair B, Wojcik SE, Valeri N, Calore F, Sampath D, Fanini F, Vannini I, Musuraca G, Dell'Aquila M, Alder H, Davuluri RV, Rassenti LZ, Negrini M, Nakamura T, Amadori D, Kay NE, Rai KR, Keating MJ, Kipps TJ, Calin GA, Croce CM (2011) Association of a microRNA/TP53 feedback circuitry with pathogenesis and outcome of B-cell chronic lymphocytic leukemia. JAMA 305(1):59–67

Faber C, Horst D, Hlubek F, Kirchner T (2011) Overexpression of Dicer predicts poor survival in colorectal cancer. Eur J Cancer 47(9):1414–1419

Fejes AP, Robertson G, Bilenky M, Varhol R, Bainbridge M, Jones SJ (2008) FindPeaks 3.1: a tool for identifying areas of enrichment from massively parallel short-read sequencing technology. Bioinformatics 24(15):1729–1730

Feng J, Liu T, Qin B, Zhang Y, Liu XS (2012) Identifying ChIP-seq enrichment using MACS. Nat Protoc 7(9):1728–1740

Fischer M (2017) Census and evaluation of p53 target genes. Oncogene 36(28):3943–3956

Fischer M, Quaas M, Steiner L, Engeland K (2016) The p53-p21-DREAM-CDE/CHR pathway regulates G2/M cell cycle genes. Nucleic Acids Res 44(1):164–174

Fischer M, Steiner L, Engeland K (2014) The transcription factor p53: not a repressor, solely an activator. Cell Cycle 13(19):3037–3058

Fornari F, Gramantieri L, Giovannini C, Veronese A, Ferracin M, Sabbioni S, Calin GA, Grazi GL, Croce CM, Tavolari S, Chieco P, Negrini M, Bolondi L (2009) MiR-122/cyclin G1 interaction modulates p53 activity and affects doxorubicin sensitivity of human hepatocarcinoma cells. Cancer Res 69(14):5761–5767

Friedman RC, Farh KK, Burge CB, Bartel DP (2009) Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 19(1):92–105

Galka-Marciniak P, Urbanek-Trzeciak MO, Nawrocka PM, Dutkiewicz A, Giefing M, Lewandowska MA, Kozlowski P (2019) Somatic mutations in miRNA genes in lung cancer-potential functional consequences of non-coding sequence variants. Cancers (Basel) 11(6):793

Garg S, Sharp PA (2016) GENE EXPRESSION. Single-cell variability guided by microRNAs Science 352(6292):1390–1391

Garibaldi F, Falcone E, Trisciuoglio D, Colombo T, Lisek K, Walerych D, Del Sal G, Paci P, Bossi G, Piaggio G, Gurtner A (2016) Mutant p53 inhibits miRNA biogenesis by interfering with the microprocessor complex. Oncogene 35(29):3760–3770

Georges SA, Biery MC, Kim SY, Schelter JM, Guo J, Chang AN, Jackson AL, Carleton MO, Linsley PS, Cleary MA, Chau BN (2008) Coordinated regulation of cell cycle transcripts by p53-inducible microRNAs, miR-192 and miR-215. Cancer Res 68(24):10105–10112

Green DR, Kroemer G (2009) Cytoplasmic functions of the tumour suppressor p53. Nature 458(7242):1127–1130

Gregorova J, Vychytilova-Faltejskova P, Sevcikova S (2021) Epigenetic regulation of MicroRNA clusters and families during tumor development. Cancers (Basel) 13(6):1333

Gregory PA, Bert AG, Paterson EL, Barry SC, Tsykin A, Farshid G, Vadas MA, Khew-Goodall Y, Goodall GJ (2008) The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. Nat Cell Biol 10(5):593–601

Hackenberg M, Rodriguez-Ezpeleta N, Aransay AM (2011) miRanalyzer: an update on the detection and analysis of microRNAs in high-throughput sequencing experiments. Nucleic Acids Res 39(Web Server issue):W132–W138

Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M Jr, Jungkamp AC, Munschauer M, Ulrich A, Wardle GS, Dewell S, Zavolan M, Tuschl T (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141(1):129–141

Hahn S, Jackstadt R, Siemens H, Hunten S, Hermeking H (2013) SNAIL and miR-34a feed-forward regulation of ZNF281/ZBP99 promotes epithelial-mesenchymal transition. EMBO J 32(23): 3079–3095

Harrison CJ, Mazzullo H, Cheung KL, Gerrard G, Jalali GR, Mehta A, Osier DG, Orchard KH (2003) Cytogenetics of multiple myeloma: interpretation of fluorescence in situ hybridization results. Br J Haematol 120(6):944–952

He L, He X, Lim LP, de Stanchina E, Xuan Z, Liang Y, Xue W, Zender L, Magnus J, Ridzon D, Jackson AL, Linsley PS, Chen C, Lowe SW, Cleary MA, Hannon GJ (2007a) A microRNA component of the p53 tumour suppressor network. Nature 447(7148):1130–1134

He X, He L, Hannon GJ (2007b) The guardian's little helper: microRNAs in the p53 tumor suppressor network. Cancer Res 67(23):11099–11101

Hendrickson DG, Hogan DJ, Herschlag D, Ferrell JE, Brown PO (2008) Systematic identification of mRNAs recruited to argonaute 2 by specific microRNAs and corresponding changes in transcript abundance. PLoS One 3(5):e2126

Hermeking H (2003) The 14-3-3 cancer connection. Nat Rev Cancer 3(12):931–943

Hermeking H (2007) p53 enters the microRNA world. Cancer Cell 12(5):414–418

Hermeking H (2010) The miR-34 family in cancer and apoptosis. Cell Death Differ 17(2):193–199

Hermeking H (2012) MicroRNAs in the p53 network: micromanagement of tumour suppression. Nat Rev Cancer 12(9):613–626

Herrera-Merchan A, Cerrato C, Luengo G, Dominguez O, Piris MA, Serrano M, Gonzalez S (2010) miR-33-mediated downregulation of p53 controls hematopoietic stem cell self-renewal. Cell Cycle 9(16):3277–3285

Hill DA, Ivanovich J, Priest JR, Gurnett CA, Dehner LP, Desruisseau D, Jarzembowski JA, Wikenheiser-Brokamp KA, Suarez BK, Whelan AJ, Williams G, Bracamontes D, Messinger Y, Goodfellow PJ (2009) DICER1 mutations in familial pleuropulmonary blastoma. Science 325(5943):965

Hoffman Y, Pilpel Y, Oren M (2014) microRNAs and Alu elements in the p53-Mdm2-Mdm4 regulatory network. J Mol Cell Biol 6(3):192–197

Hoh J, Jin S, Parrado T, Edington J, Levine AJ, Ott J (2002) The p53MH algorithm and its application in detecting p53-responsive genes. Proc Natl Acad Sci U S A 99(13):8467–8472

Hu W, Chan CS, Wu R, Zhang C, Sun Y, Song JS, Tang LH, Levine AJ, Feng Z (2010) Negative regulation of tumor suppressor p53 by microRNA miR-504. Mol Cell 38(5):689–699

Hu X, Macdonald DM, Huettner PC, Feng Z, El Naqa IM, Schwarz JK, Mutch DG, Grigsby PW, Powell SN, Wang X (2009) A miR-200 microRNA cluster as prognostic marker in advanced ovarian cancer. Gynecol Oncol 114(3):457–464

Hu Y, Lan W, Miller D (2017) Next-generation sequencing for MicroRNA expression profile. Methods Mol Biol 1617:169–177

Hwang CI, Matoso A, Corney DC, Flesken-Nikitin A, Korner S, Wang W, Boccaccio C, Thorgeirsson SS, Comoglio PM, Hermeking H, Nikitin AY (2011) Wild-type p53 controls cell motility and invasion by dual regulation of MET expression. Proc Natl Acad Sci U S A 108(34):14240–14245

Kahlert C, Klupp F, Brand K, Lasitschka F, Diederichs S, Kirchberg J, Rahbari N, Dutta S, Bork U, Fritzmann J, Reissfelder C, Koch M, Weitz J (2011) Invasion front-specific expression and prognostic significance of microRNA in colorectal liver metastases. Cancer Sci 102(10): 1799–1807

Kaller M, Liffers ST, Oeljeklaus S, Kuhlmann K, Roh S, Hoffmann R, Warscheid B, Hermeking H (2011) Genome-wide characterization of miR-34a induced changes in protein and mRNA expression by a combined pulsed SILAC and microarray analysis. Mol Cell Proteomics 10(8):M111 010462

Karaayvaz M, Pal T, Song B, Zhang C, Georgakopoulos P, Mehmood S, Burke S, Shroyer K, Ju J (2011) Prognostic significance of miR-215 in colon cancer. Clin Colorectal Cancer 10(4): 340–347

Karginov FV, Conaco C, Xuan Z, Schmidt BH, Parker JS, Mandel G, Hannon GJ (2007) A biochemical approach to identifying microRNA targets. Proc Natl Acad Sci U S A 104(49): 19291–19296

Karube Y, Tanaka H, Osada H, Tomida S, Tatematsu Y, Yanagisawa K, Yatabe Y, Takamizawa J, Miyoshi S, Mitsudomi T, Takahashi T (2005) Reduced expression of Dicer associated with poor prognosis in lung cancer patients. Cancer Sci 96(2):111–115

Kastenhuber ER, Lowe SW (2017) Putting p53 in context. Cell 170(6):1062–1078

Kaur S, Lotsari-Salomaa JE, Seppanen-Kaijansinkko R, Peltomaki P (2016) MicroRNA methylation in colorectal cancer. Adv Exp Med Biol 937:109–122

Keck T, Brabletz T (2011) Under stress: p53 controls EMT and stemness in pancreatic epithelial cells. Cell Cycle 10(11):1715

Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E (2007) The role of site accessibility in microRNA target recognition. Nat Genet 39(10):1278–1284

Kim D, Langmead B, Salzberg SL (2015) HISAT: a fast spliced aligner with low memory requirements. Nat Methods 12(4):357–360

Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL (2013) TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome Biol 14(4):R36

Kim NH, Kim HS, Kim NG, Lee I, Choi HS, Li XY, Kang SE, Cha SY, Ryu JK, Na JM, Park C, Kim K, Lee S, Gumbiner BM, Yook JI, Weiss SJ (2011a) p53 and MicroRNA-34 are suppressors of canonical Wnt signaling. Sci Signal 4(197):ra71

Kim NH, Kim HS, Li XY, Lee I, Choi HS, Kang SE, Cha SY, Ryu JK, Yoon D, Fearon ER, Rowe RG, Lee S, Maher CA, Weiss SJ, Yook JI (2011b) A p53/miRNA-34 axis regulates Snail1-dependent cancer cell epithelial-mesenchymal transition. J Cell Biol 195(3):417–433

Kim T, Veronese A, Pichiorri F, Lee TJ, Jeon YJ, Volinia S, Pineau P, Marchio A, Palatini J, Suh SS, Alder H, Liu CG, Dejean A, Croce CM (2011c) p53 regulates epithelial-mesenchymal transition through microRNAs targeting ZEB1 and ZEB2. J Exp Med 208(5):875–883

Klein U, Lia M, Crespo M, Siegel R, Shen Q, Mo T, Ambesi-Impiombato A, Califano A, Migliazza A, Bhagat G, Dalla-Favera R (2010) The DLEU2/miR-15a/16-1 cluster controls B cell proliferation and its deletion leads to chronic lymphocytic leukemia. Cancer Cell 17(1): 28–40

Kohlhammer H, Schwaenen C, Wessendorf S, Holzmann K, Kestler HA, Kienle D, Barth TF, Moller P, Ott G, Kalla J, Radlwimmer B, Pscherer A, Stilgenbauer S, Dohner H, Lichter P, Bentz M (2004) Genomic DNA-chip hybridization in t(11;14)-positive mantle cell lymphomas shows a high frequency of aberrations and allows a refined characterization of consensus regions. Blood 104(3):795–801

Krek A, Grun D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M, Rajewsky N (2005) Combinatorial microRNA target predictions. Nat Genet 37(5):495–500

Krell J, Stebbing J, Carissimi C, Dabrowska AF, de Giorgio A, Frampton AE, Harding V, Fulci V, Macino G, Colombo T, Castellano L (2016) TP53 regulates miRNA association with AGO2 to remodel the miRNA-mRNA interaction network. Genome Res 26(3):331–341

Krizhanovsky V, Lowe SW (2009) Stem cells: the promises and perils of p53. Nature 460(7259): 1085–1086

Kruse JP, Gu W (2009) Modes of p53 regulation. Cell 137(4):609–622

Kubo T, Toyooka S, Tsukuda K, Sakaguchi M, Fukazawa T, Soh J, Asano H, Ueno T, Muraoka T, Yamamoto H, Nasu Y, Kishimoto T, Pass HI, Matsui H, Huh NH, Miyoshi S (2011) Epigenetic silencing of microRNA-34b/c plays an important role in the pathogenesis of malignant pleural mesothelioma. Clin Cancer Res 17(15):4965–4974

Kumar M, Lu Z, Takwi AA, Chen W, Callander NS, Ramos KS, Young KH, Li Y (2011) Negative regulation of the tumor suppressor p53 gene by microRNAs. Oncogene 30(7):843–853

Kumar MS, Pester RE, Chen CY, Lane K, Chin C, Lu J, Kirsch DG, Golub TR, Jacks T (2009) Dicer1 functions as a haploinsufficient tumor suppressor. Genes Dev 23(23):2700–2704

Kumar RM, Cahan P, Shalek AK, Satija R, Daley Keyser A, Li H, Zhang J, Pardee K, Gennert D, Trombetta JJ, Ferrante TC, Regev A, Daley GQ, Collins JJ (2014) Deconstructing transcriptional heterogeneity in pluripotent stem cells. Nature 516(7529):56–61

La T, Liu GZ, Farrelly M, Cole N, Feng YC, Zhang YY, Sherwin SK, Yari H, Tabatabaee H, Yan XG, Guo ST, Liu T, Thorne RF, Jin L, Zhang XD (2018) A p53-responsive miRNA network promotes cancer cell quiescence. Cancer Res 78(23):6666–6679

Lagana A, Forte S, Giudice A, Arena MR, Puglisi PL, Giugno R, Pulvirenti A, Shasha D, Ferro A (2009) miRo: a miRNA knowledge base. Database (Oxford) 2009:bap008

Lal A, Thomas MP, Altschuler G, Navarro F, O'Day E, Li XL, Concepcion C, Han YC, Thiery J, Rajani DK, Deutsch A, Hofmann O, Ventura A, Hide W, Lieberman J (2011) Capture of microRNA-bound mRNAs identifies the tumor suppressor miR-34a as a regulator of growth factor signaling. PLoS Genet 7(11):e1002363

Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol 10(3):R25

Lataniotis L, Albrecht A, Kok FO, Monfries CAL, Benedetti L, Lawson ND, Hughes SM, Steinhofel K, Mayr M, Zampetaki A (2017) CRISPR/Cas9 editing reveals novel mechanisms of clustered microRNA regulation and function. Sci Rep 7(1):8585

Laudato S, Patil N, Abba ML, Leupold JH, Benner A, Gaiser T, Marx A, Allgayer H (2017) P53-induced miR-30e-5p inhibits colorectal cancer invasion and metastasis by targeting ITGA6 and ITGB1. Int J Cancer 141(9):1879–1890

Law CW, Chen Y, Shi W, Smyth GK (2014) Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. Genome Biol 15(2):R29

Leaderer D, Hoffman AE, Zheng T, Fu A, Weidhaas J, Paranjape T, Zhu Y (2011) Genetic and epigenetic association studies suggest a role of microRNA biogenesis gene exportin-5 (XPO5) in breast tumorigenesis. Int J Mol Epidemiol Genet 2(1):9–18

Lee KH, Lotterman C, Karikari C, Omura N, Feldmann G, Habbe N, Goggins MG, Mendell JT, Maitra A (2009) Epigenetic silencing of MicroRNA miR-107 regulates cyclin-dependent kinase 6 expression in pancreatic cancer. Pancreatology 9(3):293–301

Leite KR, Tomiyama A, Reis ST, Sousa-Canavez JM, Sanudo A, Camara-Lopes LH, Srougi M (2011) MicroRNA expression profiles in the progression of prostate cancer-from high-grade prostate intraepithelial neoplasia to metastasis. Urol Oncol

Leucci E, Cocco M, Onnis A, De Falco G, van Cleef P, Bellan C, van Rijk A, Nyagol J, Byakika B, Lazzi S, Tosi P, van Krieken H, Leoncini L (2008) MYC translocation-negative classical Burkitt lymphoma cases: an alternative pathogenetic mechanism involving miRNA deregulation. J Pathol 216(4):440–450

Leung AK, Sharp PA (2007) microRNAs: a safeguard against turmoil? Cell 130(4):581–585

Leung AK, Sharp PA (2010) MicroRNA functions in stress responses. Mol Cell 40(2):205–215

Li H, Durbin R (2009) Fast and accurate short read alignment with burrows-Wheeler transform. Bioinformatics 25(14):1754–1760

Li H, Rokavec M, Jiang L, Horst D, Hermeking H (2017) Antagonistic effects of p53 and HIF1A on microRNA-34a regulation of PPP1R11 and STAT3 and Hypoxia-induced epithelial to mesenchymal transition in colorectal cancer cells. Gastroenterology 153(2):505–520

Liberzon A (2014) A description of the molecular signatures database (MSigDB) web site. Methods Mol Biol 1150:153–160

Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P (2015) The molecular signatures database (MSigDB) hallmark gene set collection. Cell Syst 1(6):417–425

Liu C, Kelnar K, Liu B, Chen X, Calhoun-Davis T, Li H, Patrawala L, Yan H, Jeter C, Honorio S, Wiggins JF, Bader AG, Fagin R, Brown D, Tang DG (2011) The microRNA miR-34a inhibits prostate cancer stem cells and metastasis by directly repressing CD44. Nat Med 17(2):211–215

Liu J, Zhang C, Zhao Y, Feng Z (2017) MicroRNA control of p53. J Cell Biochem 118(1):7–14

Liu P (2018) MicroRNA expression analysis: next-generation sequencing. Methods Mol Biol 1783: 171–183

Lize M, Klimke A, Dobbelstein M (2011) MicroRNA-449 in cell fate determination. Cell Cycle 10(17):2874–2882

Lodygin D, Tarasov V, Epanchintsev A, Berking C, Knyazeva T, Korner H, Knyazev P, Diebold J, Hermeking H (2008) Inactivation of miR-34a by aberrant CpG methylation in multiple types of cancer. Cell Cycle 7(16):2591–2600

Loganantharaj R, Randall TA (2017) The limitations of existing approaches in improving MicroRNA target prediction accuracy. Methods Mol Biol 1617:133–158

Love MI, Huber W, Anders S (2014) Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 15(12):550

Lu Y, Baras AS, Halushka MK (2018) miRge 2.0 for comprehensive analysis of microRNA sequencing data. BMC Bioinformatics 19(1):275

Lujambio A, Calin GA, Villanueva A, Ropero S, Sanchez-Cespedes M, Blanco D, Montuenga LM, Rossi S, Nicoloso MS, Faller WJ, Gallagher WM, Eccles SA, Croce CM, Esteller M (2008) A microRNA DNA methylation signature for human cancer metastasis. Proc Natl Acad Sci U S A 105(36):13556–13561

Mackowiak SD (2011) Identification of novel and known miRNAs in deep-sequencing data with miRDeep2. Curr Protoc Bioinformatics Chapter 12:Unit 12 10

Martello G, Rosato A, Ferrari F, Manfrin A, Cordenonsi M, Dupont S, Enzo E, Guzzardo V, Rondina M, Spruce T, Parenti AR, Daidone MG, Bicciato S, Piccolo S (2010) A MicroRNA targeting dicer for metastasis control. Cell 141(7):1195–1207

Martin MG, Payton JE, Link DC (2009) Dicer and outcomes in patients with acute myeloid leukemia (AML). Leuk Res 33(8):e127

Marx V (2019) A dream of single-cell proteomics. Nat Methods 16(9):809–812

Mayr C, Bartel DP (2009) Widespread shortening of 3'UTRs by alternative cleavage and polyadenylation activates oncogenes in cancer cells. Cell 138(4):673–684

Mayr C, Hemann MT, Bartel DP (2007) Disrupting the pairing between let-7 and Hmga2 enhances oncogenic transformation. Science 315(5818):1576–1579

Melo SA, Moutinho C, Ropero S, Calin GA, Rossi S, Spizzo R, Fernandez AF, Davalos V, Villanueva A, Montoya G, Yamamoto H, Schwartz S Jr, Esteller M (2010) A genetic defect in exportin-5 traps precursor microRNAs in the nucleus of cancer cells. Cancer Cell 18(4): 303–315

Menendez D, Inga A, Resnick MA (2009) The expanding universe of p53 targets. Nat Rev Cancer 9(10):724–737

Merritt WM, Lin YG, Han LY, Kamat AA, Spannuth WA, Schmandt R, Urbauer D, Pennacchio LA, Cheng JF, Nick AM, Deavers MT, Mourad-Zeidan A, Wang H, Mueller P, Lenburg ME, Gray JW, Mok S, Birrer MJ, Lopez-Berestein G, Coleman RL, Bar-Eli M, Sood AK (2008) Dicer, Drosha, and outcomes in patients with ovarian cancer. N Engl J Med 359(25):2641–2650

Migliore C, Petrelli A, Ghiso E, Corso S, Capparuccia L, Eramo A, Comoglio PM, Giordano S (2008) MicroRNAs impair MET-mediated invasive growth. Cancer Res 68(24):10128–10136

Miranda KC, Huynh T, Tay Y, Ang YS, Tam WL, Thomson AM, Lim B, Rigoutsos I (2006) A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell 126(6):1203–1217

Morales S, Monzo M, Navarro A (2017) Epigenetic regulation mechanisms of microRNA expression. Biomol Concepts 8(5-6):203–212

Mudhasani R, Zhu Z, Hutvagner G, Eischen CM, Lyle S, Hall LL, Lawrence JB, Imbalzano AN, Jones SN (2008) Loss of miRNA biogenesis induces p19Arf-p53 signaling and senescence in primary cells. J Cell Biol 181(7):1055–1063

Musumeci M, Coppola V, Addario A, Patrizii M, Maugeri-Sacca M, Memeo L, Colarossi C, Francescangeli F, Biffoni M, Collura D, Giacobbe A, D'Urso L, Falchi M, Venneri MA, Muto G, De Maria R, Bonci D (2011) Control of tumor and microenvironment cross-talk by miR-15a and miR-16 in prostate cancer. Oncogene 30(41):4231–4242

Neves R, Scheel C, Weinhold S, Honisch E, Iwaniuk KM, Trompeter HI, Niederacher D, Wernet P, Santourlidis S, Uhrberg M (2010) Role of DNA methylation in miR-200c/141 cluster silencing in invasive breast cancer cells. BMC Res Notes 3:219

Nishida N, Yokobori T, Mimori K, Sudo T, Tanaka F, Shibata K, Ishii H, Doki Y, Kuwano H, Mori M (2011) MicroRNA miR-125b is a prognostic marker in human colorectal cancer. Int J Oncol 38(5):1437–1443

Oner MG, Rokavec M, Kaller M, Bouznad N, Horst D, Kirchner T, Hermeking H (2018) Combined inactivation of TP53 and MIR34A promotes colorectal cancer development and progression in mice via increasing levels of IL6R and PAI1. Gastroenterology 155(6):1868–1882

Ong SE, Blagoev B, Kratchmarova I, Kristensen DB, Steen H, Pandey A, Mann M (2002) Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. Mol Cell Proteomics 1(5):376–386

Orom UA, Lund AH (2007) Isolation of microRNA targets using biotinylated synthetic microRNAs. Methods 43(2):162–165

Orom UA, Lund AH (2010) Experimental identification of microRNA targets. Gene 451(1-2):1–5

Ory B, Ellisen LW (2011) A microRNA-dependent circuit controlling p63/p73 homeostasis: p53 family cross-talk meets therapeutic opportunity. Oncotarget 2(3):259–264

Pallasch CP, Patz M, Park YJ, Hagist S, Eggle D, Claus R, Debey-Pascher S, Schulz A, Frenzel LP, Claasen J, Kutsch N, Krause G, Mayr C, Rosenwald A, Plass C, Schultze JL, Hallek M, Wendtner CM (2009) miRNA deregulation by epigenetic silencing disrupts suppression of the oncogene PLAG1 in chronic lymphocytic leukemia. Blood 114(15):3255–3264

Pichiorri F, Suh SS, Rocci A, De Luca L, Taccioli C, Santhanam R, Zhou W, Benson DM Jr, Hofmainster C, Alder H, Garofalo M, Di Leva G, Volinia S, Lin HJ, Perrotti D, Kuehl M, Aqeilan RI, Palumbo A, Croce CM (2010) Downregulation of p53-inducible microRNAs 192, 194, and 215 impairs the p53/MDM2 autoregulatory loop in multiple myeloma development. Cancer Cell 18(4):367–381

Porkka KP, Ogg EL, Saramaki OR, Vessella RL, Pukkila H, Lahdesmaki H, van Weerden WM, Wolf M, Kallioniemi OP, Jenster G, Visakorpi T (2011) The miR-15a-miR-16-1 locus is homozygously deleted in a subset of prostate cancers. Genes Chromosomes Cancer 50(7): 499–509

Prosser HM, Koike-Yusa H, Cooper JD, Law FC, Bradley A (2011) A resource of vectors and ES cells for targeted deletion of microRNAs in mice. Nat Biotechnol 29(9):840–845

Rakheja D, Chen KS, Liu Y, Shukla AA, Schmid V, Chang TC, Khokhar S, Wickiser JE, Karandikar NJ, Malter JS, Mendell JT, Amatruda JF (2014) Somatic mutations in DROSHA and DICER1 impair microRNA biogenesis through distinct mechanisms in Wilms tumours. Nat Commun 2:4802

Raver-Shapira N, Marciano E, Meiri E, Spector Y, Rosenfeld N, Moskovits N, Bentwich Z, Oren M (2007) Transcriptional activation of miR-34a contributes to p53-mediated apoptosis. Mol Cell 26(5):731–743

Rehmsmeier M, Steffen P, Hochsmann M, Giegerich R (2004) Fast and effective prediction of microRNA/target duplexes. RNA 10(10):1507–1517

Riley T, Sontag E, Chen P, Levine A (2008) Transcriptional control of human p53-regulated genes. Nat Rev Mol Cell Biol 9(5):402–412

Roberts JT, Borchert GM (2017) Computational prediction of MicroRNA target genes, target prediction databases, and web resources. Methods Mol Biol 1617:109–122

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26(1):139–140

Rokavec M, Li H, Jiang L, Hermeking H (2014a) The p53/miR-34 axis in development and disease. J Mol Cell Biol 6(3):214–230

Rokavec M, Oner MG, Li H, Jackstadt R, Jiang L, Lodygin D, Kaller M, Horst D, Ziegler PK, Schwitalla S, Slotta-Huspenina J, Bader FG, Greten FR, Hermeking H (2014b) IL-6R/STAT3/miR-34a feedback loop promotes EMT-mediated colorectal cancer invasion and metastasis. J Clin Invest 124(4):1853–1867

Rokavec M, Bouznad N, Hermeking H (2019) Paracrine induction of epithelial-mesenchymal transition between colorectal cancer cells and its suppression by a p53/miR-192/215/NID1 axis. Cell Mol Gastroenterol Hepatol 7(4):783–802

Roldo C, Missiaglia E, Hagan JP, Falconi M, Capelli P, Bersani S, Calin GA, Volinia S, Liu CG, Scarpa A, Croce CM (2006) MicroRNA expression abnormalities in pancreatic endocrine and acinar tumors are associated with distinctive pathologic features and clinical behavior. J Clin Oncol 24(29):4677–4684

Rubio-Somoza I, Weigel D, Franco-Zorilla JM, Garcia JA, Paz-Ares J (2011) ceRNAs: miRNA target mimic mimics. Cell 147(7):1431–1432

Sachdeva M, Zhu S, Wu F, Wu H, Walia V, Kumar S, Elble R, Watabe K, Mo YY (2009) p53 represses c-Myc through induction of the tumor suppressor miR-145. Proc Natl Acad Sci U S A 106(9):3207–3212

Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP (2011) A ceRNA hypothesis: the Rosetta stone of a hidden RNA language? Cell 146(3):353–358

Schmiedel JM, Klemm SL, Zheng Y, Sahay A, Bluthgen N, Marks DS, van Oudenaarden A (2015) Gene expression. MicroRNA control of protein expression noise. Science 348(6230):128–132

Schubert J, Brabletz T (2011) p53 spreads out further: suppression of EMT and stemness by activating miR-200c expression. Cell Res 21(5):705–707

Selbach M, Schwanhausser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N (2008) Widespread changes in protein synthesis induced by microRNAs. Nature 455(7209):58–63

Shi L, Jackstadt R, Siemens H, Li H, Kirchner T, Hermeking H (2014) p53-induced miR-15a/16-1 and AP4 form a double-negative feedback loop to regulate epithelial-mesenchymal transition and metastasis in colorectal cancer. Cancer Res 74(2):532–542

Shi X, Kaller M, Rokavec M, Kirchner T, Horst D, Hermeking H (2020) Characterization of a p53/miR-34a/CSF1R/STAT3 feedback loop in colorectal cancer. Cell Mol Gastroenterol Hepatol 10(2):391–418

Shimono Y, Zabala M, Cho RW, Lobo N, Dalerba P, Qian D, Diehn M, Liu H, Panula SP, Chiao E, Dirbas FM, Somlo G, Pera RA, Lao K, Clarke MF (2009) Downregulation of miRNA-200c links breast cancer stem cells with normal stem cells. Cell 138(3):592–603

Shinozaki A, Sakatani T, Ushiku T, Hino R, Isogai M, Ishikawa S, Uozaki H, Takada K, Fukayama M (2010) Downregulation of microRNA-200 in EBV-associated gastric carcinoma. Cancer Res 70(11):4719–4727

Siemens H, Jackstadt R, Hunten S, Kaller M, Menssen A, Gotz U, Hermeking H (2011) miR-34 and SNAIL form a double-negative feedback loop to regulate epithelial-mesenchymal transitions. Cell Cycle 10(24)

Smeenk L, van Heeringen SJ, Koeppel M, van Driel MA, Bartels SJ, Akkers RC, Denissov S, Stunnenberg HG, Lohrum M (2008) Characterization of genome-wide p53-binding sites upon stress response. Nucleic Acids Res 36(11):3639–3654

Snuderl M, Kannan K, Pfaff E, Wang S, Stafford JM, Serrano J, Heguy A, Ray K, Faustin A, Aminova O, Dolgalev I, Stapleton SL, Zagzag D, Chiriboga L, Gardner SL, Wisoff JH, Golfinos JG, Capper D, Hovestadt V, Rosenblum MK, Placantonakis DG, LeBoeuf SE, Papagiannakopoulos TY, Chavez L, Ahsan S, Eberhart CG, Pfister SM, Jones DTW, Karajannis MA (2018) Recurrent homozygous deletion of DROSHA and microduplication of PDE4DIP in pineoblastoma. Nat Commun 9(1):2868

Soussi T (2011) TP53 mutations in human cancer: database reassessment and prospects for the next decade. Adv Cancer Res 110:107–139

Sticht C, De La Torre C, Parveen A, Gretz N (2018) miRWalk: an online resource for prediction of microRNA binding sites. PLoS One 13(10):e0206239

Stilgenbauer S, Nickolenko J, Wilhelm J, Wolf S, Weitz S, Dohner K, Boehm T, Dohner H, Lichter P (1998) Expressed sequences as candidates for a novel tumor suppressor gene at band 13q14 in B-cell chronic lymphocytic leukemia and mantle cell lymphoma. Oncogene 16(14):1891–1897

Su X, Chakravarti D, Cho MS, Liu L, Gi YJ, Lin YL, Leung ML, El-Naggar A, Creighton CJ, Suraokar MB, Wistuba I, Flores ER (2010) TAp63 suppresses metastasis through coordinate regulation of Dicer and miRNAs. Nature 467(7318):986–990

Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP (2005) Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 102(43):15545–15550

Suh SO, Chen Y, Zaman MS, Hirata H, Yamamura S, Shahryari V, Liu J, Tabatabai ZL, Kakar S, Deng G, Tanaka Y, Dahiya R (2011) MicroRNA-145 is regulated by DNA methylation and p53 gene mutation in prostate cancer. Carcinogenesis 32(5):772–778

Suzuki H, Yamamoto E, Nojima M, Kai M, Yamano HO, Yoshikawa K, Kimura T, Kudo T, Harada E, Sugai T, Takamaru H, Niinuma T, Maruyama R, Yamamoto H, Tokino T, Imai K, Toyota M, Shinomura Y (2010) Methylation-associated silencing of microRNA-34b/c in gastric cancer and its involvement in an epigenetic field defect. Carcinogenesis 31(12):2066–2073

Suzuki HI, Yamagata K, Sugimoto K, Iwamoto T, Kato S, Miyazono K (2009) Modulation of microRNA processing by p53. Nature 460(7254):529–533

Swarbrick A, Woods SL, Shaw A, Balakrishnan A, Phua Y, Nguyen A, Chanthery Y, Lim L, Ashton LJ, Judson RL, Huskey N, Blelloch R, Haber M, Norris MD, Lengyel P, Hackett CS, Preiss T, Chetcuti A, Sullivan CS, Marcusson EG, Weiss W, L'Etoile N, Goga A (2010) miR-380-5p represses p53 to control cellular survival and is associated with poor outcome in MYCN-amplified neuroblastoma. Nat Med 16(10):1134–1140

Tarasov V, Jung P, Verdoodt B, Lodygin D, Epanchintsev A, Menssen A, Meister G, Hermeking H (2007) Differential regulation of microRNAs by p53 revealed by massively parallel sequencing: miR-34a is a p53 target that induces apoptosis and G1-arrest. Cell Cycle 6(13):1586–1593

Tay Y, Kats L, Salmena L, Weiss D, Tan SM, Ala U, Karreth F, Poliseno L, Provero P, Di Cunto F, Lieberman J, Rigoutsos I, Pandolfi PP (2011) Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. Cell 147(2):344–357

Tay Y, Tan SM, Karreth FA, Lieberman J, Pandolfi PP (2014) Characterization of dual PTEN and p53-targeting microRNAs identifies microRNA-638/Dnm2 as a two-hit oncogenic locus. Cell Rep 8(3):714–722

Tazawa H, Tsuchiya N, Izumiya M, Nakagama H (2007) Tumor-suppressive miR-34a induces senescence-like growth arrest through modulation of the E2F pathway in human colon cancer cells. Proc Natl Acad Sci U S A 104(39):15472–15477

Tellez CS, Juri DE, Do K, Bernauer AM, Thomas CL, Damiani LA, Tessema M, Leng S, Belinsky SA (2011) EMT and stem cell-like properties associated with miR-205 and miR-200 epigenetic silencing are early manifestations during carcinogen-induced transformation of human lung epithelial cells. Cancer Res 71(8):3087–3097

Thorstensen L, Qvist H, Heim S, Liefers GJ, Nesland JM, Giercksky KE, Lothe RA (2000) Evaluation of 1p losses in primary carcinomas, local recurrences and peripheral metastases from colorectal cancer patients. Neoplasia 2(6):514–522

Tian S, Huang S, Wu S, Guo W, Li J, He X (2010) MicroRNA-1285 inhibits the expression of p53 by directly targeting its 3′ untranslated region. Biochem Biophys Res Commun 396(2):435–439

Torrezan GT, Ferreira EN, Nakahata AM, Barros BD, Castro MT, Correa BR, Krepischi AC, Olivieri EH, Cunha IW, Tabori U, Grundy PE, Costa CM, de Camargo B, Galante PA, Carraro DM (2014) Recurrent somatic mutation in DROSHA induces microRNA profile changes in Wilms tumour. Nat Commun 5:4039

Toyota M, Suzuki H, Sasaki Y, Maruyama R, Imai K, Shinomura Y, Tokino T (2008) Epigenetic silencing of microRNA-34b/c and B-cell translocation gene 4 is associated with CpG island methylation in colorectal cancer. Cancer Res 68(11):4123–4132

Trapnell C, Roberts A, Goff L, Pertea G, Kim D, Kelley DR, Pimentel H, Salzberg SL, Rinn JL, Pachter L (2012) Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. Nat Protoc 7(3):562–578

Velculescu VE, Zhang L, Vogelstein B, Kinzler KW (1995) Serial analysis of gene expression. Science 270(5235):484–487

Vistain LF, Tay S (2021) Single-cell proteomics. Trends Biochem Sci:S0968-0004(0921)00027-X.

Vogelstein B, Lane D, Levine AJ (2000) Surfing the p53 network. Nature 408(6810):307–310

Vogt M, Munding J, Gruner M, Liffers ST, Verdoodt B, Hauk J, Steinstraesser L, Tannapfel A, Hermeking H (2011) Frequent concomitant inactivation of miR-34a and miR-34b/c by CpG methylation in colorectal, pancreatic, mammary, ovarian, urothelial, and renal cell carcinomas and soft tissue sarcomas. Virchows Arch 458:313–322

Vousden KH, Prives C (2009) Blinded by the light: the growing complexity of p53. Cell 137(3): 413–431

Vousden KH, Ryan KM (2009) p53 and metabolism. Nat Rev Cancer 9(10):691–700

Wada M, Okamura T, Okada M, Teramura M, Masuda M, Motoji T, Mizoguchi H (1999) Frequent chromosome arm 13q deletion in aggressive non-Hodgkin's lymphoma. Leukemia 13(5): 792–798

Walz AL, Ooms A, Gadd S, Gerhard DS, Smith MA, Guidry Auvil JM, Meerzaman D, Chen QR, Hsu CH, Yan C, Nguyen C, Hu Y, Bowlby R, Brooks D, Ma Y, Mungall AJ, Moore RA, Schein J, Marra MA, Huff V, Dome JS, Chi YY, Mullighan CG, Ma J, Wheeler DA, Hampton OA, Jafari N, Ross N, Gastier-Foster JM, Perlman EJ (2015) Recurrent DGCR8, DROSHA, and SIX homeodomain mutations in favorable histology Wilms tumors. Cancer Cell 27(2):286–297

Wang N, Zheng J, Chen Z, Liu Y, Dura B, Kwak M, Xavier-Ferrucio J, Lu YC, Zhang M, Roden C, Cheng J, Krause DS, Ding Y, Fan R, Lu J (2019) Single-cell microRNA-mRNA co-sequencing reveals non-genetic heterogeneity and mechanisms of microRNA regulation. Nat Commun 10(1):95

Wang Z, Chen Z, Gao Y, Li N, Li B, Tan F, Tan X, Lu N, Sun Y, Sun J, Sun N, He J (2011) DNA hypermethylation of microRNA-34b/c has prognostic value for stage non-small cell lung cancer. Cancer Biol Ther 11(5):490–496

Wegert J, Ishaque N, Vardapour R, Georg C, Gu Z, Bieg M, Ziegler B, Bausenwein S, Nourkami N, Ludwig N, Keller A, Grimm C, Kneitz S, Williams RD, Chagtai T, Pritchard-Jones K, van Sluis P, Volckmann R, Koster J, Versteeg R, Acha T, O'Sullivan MJ, Bode PK, Niggli F, Tytgat GA, van Tinteren H, van den Heuvel-Eibrink MM, Meese E, Vokuhl C, Leuschner I, Graf N, Eils R, Pfister SM, Kool M, Gessler M (2015) Mutations in the SIX1/2 pathway and the

DROSHA/DGCR8 miRNA microprocessor complex underlie high-risk blastemal type Wilms tumors. Cancer Cell 27(2):298–311

Wei CL, Wu Q, Vega VB, Chiu KP, Ng P, Zhang T, Shahab A, Yong HC, Fu Y, Weng Z, Liu J, Zhao XD, Chew JL, Lee YL, Kuznetsov VA, Sung WK, Miller LD, Lim B, Liu ET, Yu Q, Ng HH, Ruan Y (2006) A global map of p53 transcription-factor binding sites in the human genome. Cell 124(1):207–219

Welch C, Chen Y, Stallings RL (2007) MicroRNA-34a functions as a potential tumor suppressor by inducing apoptosis in neuroblastoma cells. Oncogene 26(34):5017–5022

Wiklund ED, Bramsen JB, Hulf T, Dyrskjot L, Ramanathan R, Hansen TB, Villadsen SB, Gao S, Ostenfeld MS, Borre M, Peter ME, Orntoft TF, Kjems J, Clark SJ (2011) Coordinated epigenetic repression of the miR-200 family and miR-205 in invasive bladder cancer. Int J Cancer 128(6):1327–1334

Wong TS, Liu XB, Wong BY, Ng RW, Yuen AP, Wei WI (2008) Mature miR-184 as potential oncogenic microRNA of squamous cell carcinoma of tongue. Clin Cancer Res 14(9):2588–2592

Xi Y, Formentini A, Chien M, Weir DB, Russo JJ, Ju J, Kornmann M (2006) Prognostic values of microRNAs in colorectal cancer. Biomark Insights 2:113–121

Xiao J, Lin H, Luo X, Wang Z (2011) miR-605 joins p53 network to form a p53:miR-605:Mdm2 positive feedback loop in response to stress. EMBO J 30(3):524–532

Xu N, Papagiannakopoulos T, Pan G, Thomson JA, Kosik KS (2009) MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. Cell 137(4):647–658

Yamakuchi M, Ferlito M, Lowenstein CJ (2008) miR-34a repression of SIRT1 regulates apoptosis. Proc Natl Acad Sci U S A 105(36):13421–13426

Yamakuchi M, Lotterman CD, Bao C, Hruban RH, Karim B, Mendell JT, Huso D, Lowenstein CJ (2010) P53-induced microRNA-107 inhibits HIF-1 and tumor angiogenesis. Proc Natl Acad Sci U S A 107(14):6334–6339

Zhang XJ, Ye H, Zeng CW, He B, Zhang H, Chen YQ (2010) Dysregulation of miR-15a and miR-214 in human pancreatic cancer. J Hematol Oncol 3:46

Zhao J, Lammers P, Torrance CJ, Bader AG (2013) TP53-independent function of miR-34a via HDAC1 and p21(CIP1/WAF1.). Mol Ther 21(9):1678–1686

Chapter 8
Machine Learning Using Gene-Sets to Infer miRNA Function

Andrew Dhawan and Francesca M. Buffa

Abstract miRNA are regulators of cell phenotype, and there is clear evidence that these small posttranscriptional modifiers of gene expression are involved in defining a cellular response across states of development and disease. Classical methods for elucidating the repressive effect of a miRNA on its targets involve controlling for the many factors influencing miRNA action, and this can be achieved in cell lines, but misses tissue and organism level context which are key to a miRNA function. Also, current technology to carry out this validation is limited in both generalizability and throughput. Methodologies with greater scalability and rapidity are required to better understand the function of these important species of RNA. To this end, there is an increasing store of RNA expression level data incorporating both miRNA and mRNA, and in this chapter, we describe how to use machine learning and gene-sets to translate the knowledge of phenotype defined by mRNA to putative roles for miRNA. We outline our approach to this process and highlight how it was done for our miRNA annotation of the hallmarks of cancer using the Cancer Genome Atlas (TCGA) dataset. The concepts we present are applicable across datasets and phenotypes, and we highlight potential pitfalls and challenges that may be faced as they are used.

Keywords miRNA · miRNA function · Regularized regression · Machine learning · Hallmarks of cancer · Gene-set · Gene signatures · Expression · miRNA–mRNA network

A. Dhawan
Neurological Institute, Cleveland Clinic Foundation, Cleveland, OH, USA

Medical Sciences Division, University of Oxford, Oxford, UK

F. M. Buffa (✉)
Medical Sciences Division, University of Oxford, Oxford, UK

Department of Computing Sciences, Bocconi University and Bocconi Institute for Data Science and Analytics, Milano, Italy
e-mail: francesca.buffa@oncology.ox.ac.uk

U. Schmitz et al. (eds.), *Systems Biology of MicroRNAs in Cancer*, Advances in Experimental Medicine and Biology 1385,
https://doi.org/10.1007/978-3-031-08356-3_8

8.1 Introduction

miRNA are short noncoding RNAs that regulate gene expression posttranscriptionally. They classically function by exerting a repressive effect on target mRNA transcripts, resulting in transcript degradation and inhibition of translation. As a result, miRNA reduces functional protein levels through RNA-induced silencing complex (RISC) machinery (Chen and Rajewsky 2007; Cloney 2016). How miRNAs are thought to target cognate transcripts is through base-pair matching on an 8-nucleotide recognition site usually in the 3′ untranslated region of a transcript. This is a thermodynamic process for which much work has already been done to understand the key variables increasing the likelihood of a given mRNA being targeted by a particular miRNA. As a result, there are multiple databases and methodologies that provide a listing of miRNA and putative mRNA targets which we will use for our analyses, reviewed in more detail in Peterson et al. (2014). These databases are each based on slightly different target prediction methods, and to capture the range of possibilities, we considered the union of the sets of mRNA predicted to be targeted by each miRNA (Oliveira et al. 2017). In doing so, we hope to have reduced any element of bias that could have been introduced into our results by anyone target prediction approach, and these are easily implemented through the use of R packages, such as miRNAtap (Pajak and Simpson 2016). We summarize the various databases and briefly describe their methodologies in Table 8.1.

As miRNAs target thousands of transcripts, and conversely one target can be actioned upon by multiple miRNA, these short noncoding RNAs potentially carry a great deal of biological information regarding cellular state. Indeed, their expression is known to regulate many of the phenotypes associated with development, oncogenesis, and therapeutic response (Aguda et al. 2008; Bartel 2004; Carroll et al. 2013; Chen et al. 2014; Gee et al. 2015). The association of miRNA to these phenotypes has typically been arduous, requiring targeted experiments demonstrating relationships between specific phenotypic assays and miRNA expression in vitro with concordant evidence of miRNA targeting relevant mRNAs. Needless to say, such well-conducted experiments are few in number, and their results are often generalized and extrapolated to in vivo results without clear validation.

Given the limited throughput of such in vitro studies, more sustainable ways to relate miRNA expression to a phenotype can be extremely helpful, and can help in refining the design of in vitro experiments. We present here an approach to translate information from better understood mRNA gene sets (or expression signatures) to infer miRNA function. The backbone of this is the gene expression signature—the quantity that relates phenotype to mRNA expression (Liberzon et al. 2015). A gene expression signature for this discussion will be defined as a set of genes, whose expression, when summarized into a single value (score) by elementary summary statistics such as mean or median, or more complex linear combinations, covaries with a phenotype of interest in a given population of samples. Generally, gene signatures defined using this paradigm are based on mRNA expression levels, and

Table 8.1 Comparison of miRNA target prediction algorithms adapted from Supplementary Information of Buffa et al. (2011)

Name	Characteristics	References
TargetScan	• Perfect seed match rule with specific seed 5′ and/or 3′ flank requirements. • Thermodynamics of seed matching considered. • Conservation among human, chimp, rodent, and dog.	Lewis et al. (2005)
PicTar	• Preference for perfect seed match. • Optimal binding free energy considered when choosing targets. • Maximum likelihood model for multiple miRNAs potentially binding a target sequence. • Conservation among human, chimp, rodent, and dog.	Lall et al. (2006)
DianaLab	• Empirically determined binding rules. • Unique thermodynamic considerations in binding based on experimental data. • Conservation among human, chimp, rodent, and dog.	Maragkakis et al. (2009)
miBridge	• Dynamic programming alignment score cutoff. • Free energy calculation for miRNA–mRNA duplex considered. • Conservation among human, rodent, and dog. • Targets containing simultaneous 5′- and 3'-UTR. • Interaction sites.	Lee et al. (2009)
miRanda	• Dynamic programing alignment score cutoff. • Free energy calculation for miRNA-mRNA duplex considered. • Conservation among human, chimp, rodent, and dog.	Miranda et al. (2006)
miRTarget2	• Seed 5′ and/or 3′ flank requirements, based on support vector machine model and large training dataset. • Free energy calculation for miRNA–mRNA duplex considered. • Conservation among human, rodent, and chicken.	Wang and El Naqa (2008)

many have been compiled, tested, and validated across cell lines and human tissues. These signatures can then be used as a "proxy" measurement of the phenotype. Databases of gene signatures can be found through many resources and primary literature, but the most well-known database has been compiled by the Broad Institute in MSigDB (Liberzon et al. 2011).

The task of summarizing gene expression values for a gene set (hence a set of gene expression values) into a single score is non-trivial, and platform-agnostic ways in which to do this that scale well with different sample sizes are not obvious. Many well-known approaches are possible, such as measures of location and variability of the distribution of gene expression values (e.g., median, mean, Z-score, and first principal component in a PCA analysis), single-sample gene-set enrichment analysis (ssGSEA), pathway-level analysis of gene expression (PLAGE) and others (Hänzelmann et al. 2013; Tomfohr et al. 2005). For the discussion in this chapter, we will focus on one of them, the median expression of signature genes (after expression normalization and batch correction, if applicable) as our summary

score, but the concepts we will present are transferable across other scores. On the other hand, in our experience, the median score tends to be easily portable across platforms, scales well with different sample sizes, and is relatively robust to outliers in gene expression. However, it can be less sensitive to pick small changes in the phenotypes when compared, for example, to a mean or Z-score. It is therefore important to evaluate the properties of a summary score, and evaluate its applicability, using standard statistical metrics. In the following subsection, we discuss the notion of gene signature quality control, a crucial step that has been missing from much of the present literature on gene signatures. Quality control is key to ensuring the appropriate use of gene signatures in more broad situations than those from which the signature was derived (Dhawan et al. 2019).

Approximately 32,284 gene sets/gene signatures are listed on MSigDB in its latest version (v.7.4), many corresponding to well-known phenotypes such as starvation, hypoxia/anoxia (low/no oxygen), angiogenesis, organogenesis, response, or resistance to stimuli (e.g., drugs and other treatments). Via a summary score, these signatures potentially "translate" mRNA expression to phenotypes, and can be used in many cases to infer, hence better understand, cellular phenotypes. Among the existing signatures with the strongest biological validation are the core "Hallmarks" gene sets, of which a subset will be used in our analyses.

Another fundamental step that informs on the miRNA function is the notion of miRNA targeting, and the expression of the miRNA targets. In our analyses, we used the Cancer Genome Atlas (TCGA), one of the most-used and well-regarded datasets in the field of cancer genomics, generated by a multicenter effort led by the National Institutes of Health in the United States (Tomczak et al. 2015). It is a compilation of genomic data with standardized clinical annotation across over 20,000 tumors from 33 cancer types. Genomic data spans whole-exome sequencing, RNA-sequencing, small RNA-sequencing, single nucleotide polymorphism (SNP) arrays, and is complemented in some cases by immunohistochemistry and proteomic analyses. This dataset has been a trove of information in cancer genomics, and due to its size has been among the strongest drivers of new information and validation in the field. Its unique nature with concurrent profiling of miRNA and mRNA expression in the same tumor samples enables a very clear analysis of how the expression of these two species relates, which will be necessary for linking miRNA to mRNA.

8.2 Statistical Preliminaries

8.2.1 Linear Modeling and Penalized Regression

The machine learning approach we will take here to determine whether a given miRNA shows anticorrelation with a given mRNA will rely on linear modeling. Briefly, linear models describe the variation in a response (output) variable, as a linear combination of the predictor (input) variables. There are a variety of methods to determine how to optimally chose parameters for the model that allows for the

linear combination of predictors to be most representative of the output, based on maximum likelihood estimates from a given dataset of known predictor and response variables. However, like any statistical model, care must be taken during the fitting of these models to assure that they are not overfitting (in essence, too tightly bound to the known dataset) or underfitting (too few known data points to make reasonable parameter estimates), and both of which limit generalizability. These issues can occur for a variety of reasons, but with linear modelling, the most common is the issue of underfitting due to too many predictor variables being considered in proportion to the size of the known dataset. For instance, with two sets of known observations, at most two parameters could be determined in a linear system, but this system would likely not generalize well. As such, because the number of observations we can use as part of our "known" dataset is fixed, we must optimize the number of predictors in the model to increase the generalizability of the linear model. This is done through parameter selection and subsequently through penalization. This is implemented through the "penalization" package in R (Goeman 2017).

Our parameter (feature) selection is done by first assessing for even weak correlations with the response variable of interest; typically, if a Spearman correlation has a statistical significance p-value less than 0.2, we consider it to be weakly correlated with the response variable, suggesting that it may be a reasonable parameter to include in the initial model. The next step in our approach is to consider L1 and L2 penalties, also known as lasso and ridge regression penalties. In doing so, we can reduce the number of parameters by shrinking certain coefficients in the linear model to zero, thereby allowing for the stronger predictors to be better represented by the coefficients of the model. For further details regarding the implementation of penalized regression, we refer the reader to (Heckman and Ramsay 2000), and for further details regarding generalized linear models and statistical modelling approaches, we refer the reader to Zheng and Agresti (2000).

8.2.2 Gene Signature Quality Control

Before using the mRNA gene expression signature in our proposed analysis, it is imperative to assure that the signature can be used reliably in the given dataset. This step is often omitted in similar analyses, and we feel that this is a significant pitfall, wherein the blind application of gene signatures can result in invalid results. We propose that a gene expression signature should only be used in the tissues and conditions for which it has been derived unless its validity has been tested in the dataset of interest.

To test signature validity, our group has proposed a quality control protocol called sigQC in which datasets and signatures are taken as input and a suite of quality control plots and metrics are produced as output. We will consider these here briefly but would refer the reader for further information to our paper on the topic (Dhawan et al. 2019). Of note, this has been implemented in an R package (available through CRAN) that has already been integrated into workflows across multiple labs.

First, to assess signature validity, one must ensure that the genes of the signature are expressed in a tissue of interest and that the genes show some variability across the samples considered. Following this, we examine the distribution of the expression of the signature genes to assure that there is not a significant skew because of a subset of the genes. The sigQC package also produces plots and metrics that test the degree of co-correlation between the genes of the signature, as depending on the metric chosen for gene signature score, the genes may need to all show positive co-correlation. Finally, sigQC also evaluates the impacts of various scoring metrics and checks the degree of co-correlation among these metrics. Metrics that are all co-correlated give higher confidence in the robustness of the gene signature score, regardless of which metric is chosen to represent the signature score. Following this, random sets of genes are sampled with these calculations repeated to derive empiric statistical distributions for each of the values produced by sigQC, thereby testing the degree of statistical significance in the choice of signature genes. For further details regarding the use or application of sigQC, we refer the reader to Dhawan et al. (2019) and https://cran.r-project.org/web/packages/sigQC/.

8.2.3 Rank Product Statistic

A key operation in our analysis is that of the rank product (Breitling et al. 2004). We utilize the rank product statistic to aggregate findings across tissue types to identify consistent associations. This is used because (i) it is non-parametric and (ii) it is easily implemented computationally (Hong et al. 2006). For a given set of variables of interest in the tissue types, the rank product statistic computes the relative rank of each of the variables in the given tissue type, then computes the product of these ranks. This product is taken as the test statistic and compared to the empiric distribution of rank products under a null model where variables of interest are randomly distributed across the tissue types. As such, this operation tests against the hypothesis that the variables of interest are randomly ordered across each of the tissues.

8.3 Approach to Choosing Representative Gene Signatures

The first step to connect mRNA expression signatures to miRNA expression was to define the key phenotypes of interest. To do this, we obtained gene expression signatures that would best represent these phenotypes in the tissues of interest. In the case of our work in Dhawan et al. (2018), our phenotypes of interest were the hallmarks of cancer, as outlined by Hanahan and Weinberg (2011, 2000). These phenotypes had the benefit of being well-described and well-known to occur in multiple cancer types, and also had well-known gene expression signatures. Also, there was a biological rationale, and a built-in “sanity check” for the analysis being

performed, as many of the hallmarks signatures, already had validated associated miRNA (e.g., miR-210 and hypoxia) (Camps et al. 2008; Gee et al. 2010). Also, for many of the hallmarks of cancer, there were multiple gene signatures, each different in their composition, which was an aspect of redundancy and robustness that we had used to our advantage in our analysis, by seeking associations strengthened by consensus. Should a given miRNA be associated with multiple signatures representative of a given phenotype, it increases the likelihood of relevance.

Among the candidate gene signatures that could be used for our analysis, we next interrogated the quality of these signature genes in the datasets of interest. In particular, we ran our *sigQC* platform across signatures and gene expression data across cancer types and determined for each signature and cancer type combination, whether its performance across the various metrics produced by the package was adequate for further analysis.

Because our computational approach was designed to use linear models to elucidate miRNA as predictors of gene signature score, the emphasis on the interpretation of the sigQC output was to assure that the gene signature score was a reasonable summary statistic for the complete set of genes and that the set of genes varied together in the same direction (either all positively or negatively in association with a hallmark). As such, the primary sigQC output metrics used for further analysis were those relating to signature gene co-correlation, signature gene expression, and the co-correlations between signature scoring metrics. First, we assured ourselves that signature genes were expressed and that the signature co-correlation ascertained whether the signature genes co-varied consistently. Following this, we tested whether the signature scoring metric chosen (in our case, the median of gene expression was used) was reliable and had a strong correlation with the alternative scoring metrics. In doing so, we were able to ensure that the signatures used were applied in a manner consistent with the datasets used.

8.4 Evidence Across Tissue Types for miRNA Associations to Key Gene Signatures

To substantiate our results, and to assure that the associations identified were not related to tissue effects alone, we sought confirmation of the miRNA associations in other tissues with similar histopathological bases (e.g., epithelial tissues). We selected preferentially for those associations between miRNA and signatures that persisted across tissue types using the rank product statistic. In this way, we established stronger evidence for the associations and filtered spurious associations that may have been tissue-specific biases. We included in this analysis the tissues that had a sufficient number of samples to warrant robust miRNA–signature associations and biologic relevance. That is, because our analysis was focused around epithelial tumors, only these tissues were included (i.e., non-epithelial tissues were excluded), to ensure that the most biologically relevant signals would be amplified.

Had this step not been done and we focused only on a single tissue type, there is the possibility that spurious associations could have emerged. On the other hand, it is also important to note that these associations by their nature, should not be presumed to be the same across similar tissues. Indeed, each tissue of the body (and even subregions of tissues) are specialized toward its biological niche and thereby has different regulatory programs. Without greater sample sizes of paired data in each tissue, we could not identify with confidence which miRNA–signature associations were tissue-specific, though it is possible that these could be identified in the future.

8.5 Hallmarks-Associated miRNA Preferentially Regulate Tumor Suppressor Genes

Up to this point, we have shown how to find the miRNA associated with phenotypes, and the next phase of our analysis is to "go backward" and ask the question of which mRNA was most strongly negatively associated with these species, to identify the candidates for mRNA–miRNA repression. In doing so, our goal is to uncover putative functional miRNA–mRNA repressive interactions driving phenotype. As noted previously, this concept is motivated by the observation that in certain cancers, several specific miRNAs (oncomiRs) are thought to be key drivers of the tumor phenotype, acting by repression of tumor suppressor genes (TSGs) (Volinia et al. 2010).

Thus, we examined the union of the set of mRNAs predicted to be targeted by the hallmarks-associated set of miRNA by using the miRNAtap R package. Correlation coefficients for the expression values of each miRNA–mRNA pair were computed. For each miRNA–mRNA pair, if the degree of anticorrelation in expression across tissue types was statistically significant using the rank product statistic, then this miRNA–mRNA pair was felt to be a potential candidate for experimental validation. We identified a statistically significant overlap with tumor suppressor genes among these mRNA, suggesting that our hypothesis that the hallmarks-associated miRNA were preferentially regulating tumor suppressor genes was correct. Null distributions were defined also with random sets of miRNA–mRNA pairs to exclude the hypothesis that these mRNAs were themselves biased in favor of tumor suppressor genes. Results are shown pictorially in Fig. 8.1.

Lastly, we endeavored to show that the miRNA–mRNA associations were statistically significant in the context of broader genetic variation across the dataset. In particular, we sought to ensure that the mRNA–mRNA association for tumor suppressor genes was not due to copy number changes (which themselves are associated with TSG expression), methylation of TSG promoters, and mutational changes (as TSG are preferentially mutated in cancers). A multivariate model was built with each of these variables, fit across tissue types, and coefficients were compared to assess whether the miRNA–mRNA interaction coefficient remained negative. Among the TSG that remained with statistically significant miRNA–

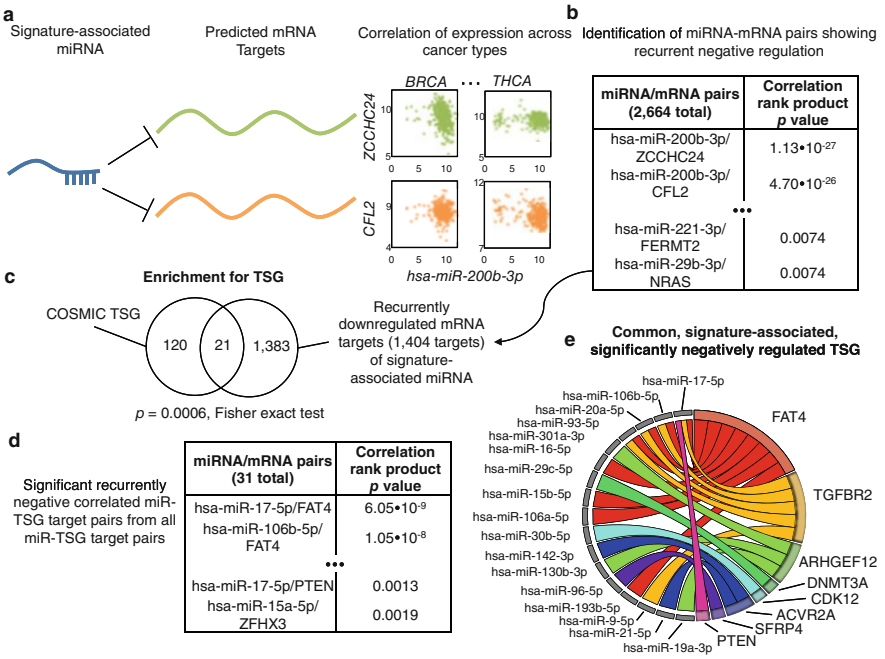


Fig. 8.1 Approach used for interpreting miRNA–target interactions. (**a**) First, miRNA–target pairs for each positively associated hallmark-associated miRNA were identified, and the correlation between these was determined. (**b**) Next, the correlations across cancer types were aggregated, and those identified as consistently negative-ranking were identified with the rank product statistic. (**c**) Among this list of miRNA–mRNA target pairs, there was highly significant enrichment for tumor suppressor genes, as identified by the Fisher exact test. (**d**) The same procedure as described in (**a**) and (**b**) was repeated for all miRNA and all predicted target TSG pairs, with each TSG considered individually. (**e**) From the lists identified in b and d, we identified those miRNA–TSG pairs in common, and plot their interactions on a circos plot, showing the repressive actions of each miRNA on its predicted target TSG. Reproduced from Dhawan et al. 2018, Nature Communications with permission. Licensed under Creative Commons Attribution 4.0: http://creativecommons.org/licenses/by/4.0/

mRNA interaction coefficients in this analysis, we showed that the samples in which there was a potential miRNA-mediated expression, there did not tend to be TSG promoter methylation, TSG mutation, or gene deletion. The predicted miRNA–TSG interactions remaining after this analysis are the most strongly predicted to be related to the hallmarks of cancer.

8.6 Conclusions and Future Directions

The effect of miRNA on cellular phenotype is not yet fully understood, despite an evolving knowledge of miRNA biology and mechanism. While the approach we have proposed describes a method for inferring functional relationships between miRNA expression and disease phenotype, the challenge of finding sufficiently large datasets remains. With smaller datasets, the impact of noise and artefactual findings is greater, and one must be mindful of this limitation. In addition to the accrual of more data, future work in this field must focus on defining novel statistical methodologies to uncover miRNA–mRNA interactions. Further, emerging evidence suggests that miRNA isoforms (isomiRs) play a role in cancer, and non-canonical miRNA–mRNA interactions may also act paradoxically in increasing mRNA and target protein levels (Stavast and Erkeland 2019; Telonis et al. 2017). As more data becomes available, our ability to see finer-grained details of these interactions will be enhanced, uncovering tissue-specific biology that could be exploited for therapeutics. While functional validation remains the gold standard, more targeted hypotheses can be formed, thereby accelerating the pace of discovery.

References

Aguda BD, Kim Y, Piper-Hunter MG, Friedman A, Marsh CB (2008) MicroRNA regulation of a cancer network: consequences of the feedback loops involving miR-17-92, E2F, and Myc. Proc Natl Acad Sci 105:19678–19683

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116:281–297

Breitling R, Armengaud P, Amtmann A, Herzyk P (2004) Rank products: a simple, yet powerful, new method to detect differentially regulated genes in replicated microarray experiments. FEBS Lett 573:83–92

Buffa FM, Camps C, Winchester L, Snell CE, Gee HE, Sheldon H, Taylor M, Harris AL, Ragoussis J (2011) microRNA associated progression pathways and potential therapeutic targets identified by integrated mRNA and microRNA expression profiling in breast cancer. Cancer Res:canres-0489

Camps C, Buffa FM, Colella S, Moore J, Sotiriou C, Sheldon H, Harris AL, Gleadle JM, Ragoussis J (2008) Hsa-miR-210 is induced by hypoxia and is an independent prognostic factor in breast cancer. Clin Cancer Res 14:1340–1348

Carroll AP, Tooney PA, Cairns MJ (2013) Context-specific microRNA function in developmental complexity. J Mol Cell Biol 5:73–84

Chen K, Rajewsky N (2007) The evolution of gene regulation by transcription factors and microRNAs. Nat Rev Genet 8:93–103. https://doi.org/10.1038/nrg1990

Chen S, Xue Y, Wu X, Le C, Bhutkar A, Bell EL, Zhang F, Langer R, Sharp PA (2014) Global microRNA depletion suppresses tumor angiogenesis. Genes Dev 28:1054–1067

Cloney R (2016) Non-coding RNA: deciphering the rules of microRNA targeting. Nat Rev Genet 17:718

Dhawan A, Barberis A, Cheng W-C, Domingo E, West C, Maughan T, Scott JG, Harris AL, Buffa FM (2019) Guidelines for using sigQC for systematic evaluation of gene signatures. Nat Protoc 14:1377–1400. https://doi.org/10.1038/s41596-019-0136-8

Dhawan A, Scott JG, Harris AL, Buffa FM (2018) Pan-cancer characterisation of microRNA across cancer hallmarks reveals microRNA-mediated downregulation of tumour suppressors. Nat Commun 9:5228. https://doi.org/10.1038/s41467-018-07657-1

Gee HE, Buffa FM, Harris AL, Toohey JM, Carroll SL, Cooper CL, Beith J, McNeil C, Carmalt H, Mak C, Warrier S, Holliday A, Selinger C, Beckers R, Kennedy C, Graham P, Swarbrick A, Millar EKA, O'Toole SA, Molloy T (2015) MicroRNA-related DNA repair/cell-cycle genes independently associated with relapse after radiation therapy for early breast cancer. Int J Radiat Oncol Biol Phys 93:1104–1114. https://doi.org/10.1016/j.ijrobp.2015.08.046

Gee HE, Camps C, Buffa FM, Patiar S, Winter SC, Betts G, Homer J, Corbridge R, Cox G, West CML, Ragoussis J, Harris AL (2010) Hsa-mir-210 is a marker of tumor hypoxia and a prognostic factor in head and neck cancer. Cancer 116:2148–2158. https://doi.org/10.1002/cncr.25009

Goeman JJ (2017) Penalized R package, version 0.9-50

Hanahan D, Weinberg RA (2011) Hallmarks of cancer: the next generation. Cell 144:646–674

Hanahan D, Weinberg RA (2000) The hallmarks of cancer. Cell 100:57–70

Hänzelmann S, Castelo R, Guinney J (2013) GSVA: gene set variation analysis for microarray and RNA-seq data. BMC Bioinformatics 14:7

Heckman NE, Ramsay JO (2000) Penalized regression with model-based penalties. Canadian J Statistic 28:241–258. https://doi.org/10.2307/3315976

Hong F, Breitling R, McEntee CW, Wittner BS, Nemhauser JL, Chory J (2006) RankProd: a bioconductor package for detecting differentially expressed genes in meta-analysis. Bioinformatics 22:2825–2827. https://doi.org/10.1093/bioinformatics/btl476

Lall S, Grün D, Krek A, Chen K, Wang Y-L, Dewey CN, Sood P, Colombo T, Bray N, MacMenamin P, others (2006) A genome-wide map of conserved microRNA targets in C. elegans. Curr Biol 16:460–471

Lee I, Ajay SS, Yook JI, Kim HS, Hong SH, Kim NH, Dhanasekaran SM, Chinnaiyan AM, Athey BD (2009) New class of microRNA targets containing simultaneous 5'-UTR and 3'-UTR interaction sites. Genome Res 19:1175–1183. https://doi.org/10.1101/gr.089367.108

Lewis BP, Burge CB, Bartel DP (2005) Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell 120:15–20

Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P (2015) The molecular signatures database hallmark gene set collection. Cell Systems 1:417–425

Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP (2011) Molecular signatures database (MSigDB) 3.0. Bioinformatics 27:1739–1740

Maragkakis M, Reczko M, Simossis VA, Alexiou P, Papadopoulos GL, Dalamagas T, Giannopoulos G, Goumas G, Koukis E, Kourtis K, others (2009) DIANA-microT web server: elucidating microRNA functions through target prediction. Nucleic Acids Res 37:W273–W276

Miranda KC, Huynh T, Tay Y, Ang Y-S, Tam W-L, Thomson AM, Lim B, Rigoutsos I (2006) A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. Cell 126:1203–1217. https://doi.org/10.1016/j.cell.2006.07.031

Oliveira AC, Bovolenta LA, Nachtigall PG, Herkenhoff ME, Lemke N, Pinhal D (2017) Combining results from distinct MicroRNA target prediction tools enhances the performance of analyses. Front Genet 8:59. https://doi.org/10.3389/fgene.2017.00059

Pajak M, Simpson TI (2016) miRNAtap: miRNAtap: microRNA targets–aggregated predictions

Peterson SM, Thompson JA, Ufkin ML, Sathyanarayana P, Liaw L, Congdon CB (2014) Common features of microRNA target prediction tools. Front Genet 5:23. https://doi.org/10.3389/fgene.2014.00023

Stavast CJ, Erkeland SJ (2019) The non-canonical aspects of MicroRNAs: many roads to gene regulation. Cell 8. https://doi.org/10.3390/cells8111465

Telonis AG, Magee R, Loher P, Chervoneva I, Londin E, Rigoutsos I (2017) Knowledge about the presence or absence of miRNA isoforms (isomiRs) can successfully discriminate amongst 32 TCGA cancer types. Nucleic Acids Res 45:2973–2985. https://doi.org/10.1093/nar/gkx082

Tomczak K, Czerwińska P, Wiznerowicz M (2015) The cancer genome atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol 19:A68

Tomfohr J, Lu J, Kepler TB (2005) Pathway level analysis of gene expression using singular value decomposition. BMC Bioinformatics 6:225

Volinia S, Galasso M, Costinean S, Tagliavini L, Gamberoni G, Drusco A, Marchesini J, Mascellani N, Sana ME, Jarour RA, others (2010) Reprogramming of miRNA networks in cancer and leukemia. Genome Res 20:589–599

Wang X, El Naqa IM (2008) Prediction of both conserved and nonconserved microRNA targets in animals. Bioinformatics 24:325–332. https://doi.org/10.1093/bioinformatics/btm595

Zheng B, Agresti A (2000) Summarizing the predictive power of a generalized linear model. Stat Med 19:1771–1781. https://doi.org/10.1002/1097-0258(20000715)19:13<1771::AID-SIM485>3.0.CO;2-P

Chapter 9
miRNA:miRNA Interactions: A Novel Mode of miRNA Regulation and Its Effect On Disease

Meredith Hill and Nham Tran

Abstract MicroRNAs (miRNAs) are known for their role in the post-transcriptional regulation of messenger RNA (mRNA). However, recent evidence has shown that miRNAs are capable of regulating non-coding RNAs, including miRNAs, in what is known as miRNA:miRNA interactions. There are three main models for the interplay between miRNAs. These involve direct interaction between two miRNAs, either in their mature or primary form, the subsequent changes in miRNA expression due to miRNA-directed transcriptional changes, and the cell-wide impact on miRNA and mRNA levels as a result of miRNA manipulation. Networks of mRNA and miRNA regulatory connections are invaluable to the discovery of miRNA:miRNA pathways, but this cannot be applied without consideration of the specific cell type or condition.

In this chapter, we discuss what is understood about miRNA:miRNA interactions, their mechanisms and consequences in disease biology, and suggest further avenues of investigation based on current gaps in the literature and in our understanding of miRNA biology. We also address the pitfalls in contemporary methods relating to the identification of miRNA:miRNA interactions. Future work in this area may ultimately change the definitional role of miRNAs, and have far-reaching impacts on therapeutic and diagnostic developments.

Keywords MicroRNA · miRNA regulation · miRNA:miRNA interaction · miRNAs in disease · miRNA networks

M. Hill
School of Biomedical Engineering, Centre for Health Technologies, Faculty of Engineering and IT, The University of Technology Sydney, NSW, Sydney, Australia

N. Tran (✉)
School of Biomedical Engineering, Centre for Health Technologies, Faculty of Engineering and IT, The University of Technology Sydney, NSW, Sydney, Australia

The Sydney Head and Neck Cancer Institute, Sydney Cancer Centre, Royal Prince Alfred Hospital, Sydney, NSW, Australia
e-mail: Nham.Tran@uts.edu.au

U. Schmitz et al. (eds.), *Systems Biology of MicroRNAs in Cancer*, Advances in Experimental Medicine and Biology 1385,
https://doi.org/10.1007/978-3-031-08356-3_9

9.1 Introduction

MicroRNAs (miRNAs) are typically known for their role in the negative regulation of messenger RNA (mRNA) via complementary binding to the 3' untranslated region (UTR). However, recent evidence suggests that miRNAs may also target non-coding RNA, including other miRNAs. This is termed a miRNA:miRNA interaction, where a miRNA influences the expression of another miRNA through direct or indirect means.

There are three main forms of miRNA:miRNA interaction. The first is those miRNA:miRNA interactions that occur through Watson-Crick pairing between either a primary miRNA (pri-miRNA) and mature miRNA, or between two mature miRNAs. In the second category, miRNAs indirectly control miRNA expression by targeting transcriptional regulators or the miRNA biogenesis components. And the third subset considers the cascading effect of miRNA:miRNA interactions on secondary mRNA and miRNA expression in the wider cellular environment (Fig. 9.1).

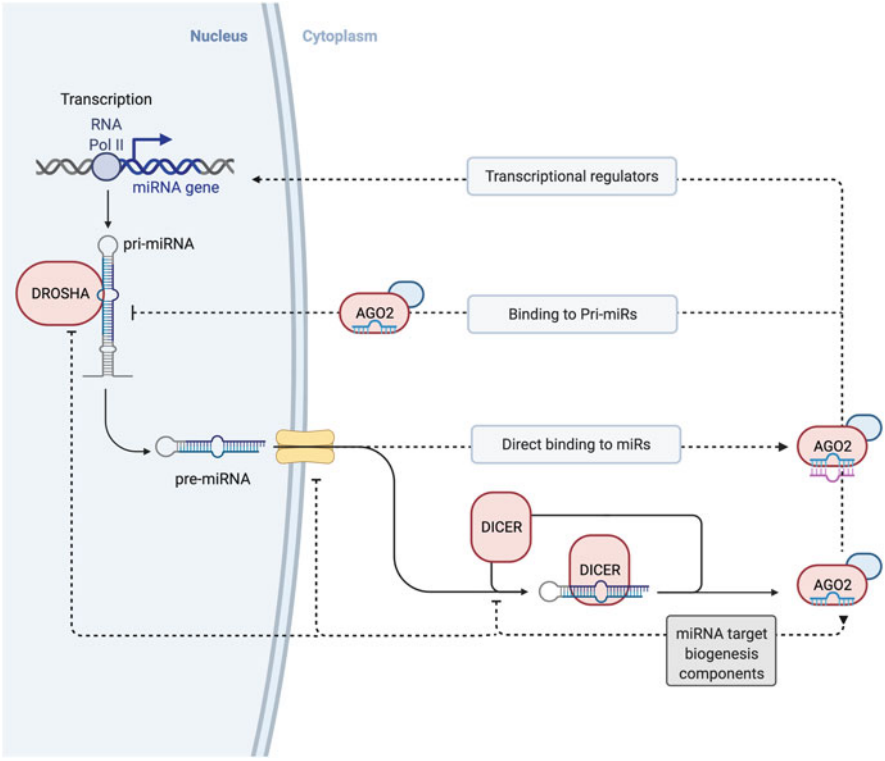


Fig. 9.1 Overview of the different forms of miRNA:miRNA interaction

This chapter will present these three different forms of miRNA:miRNA interactions, their mechanisms, their role in disease development, and current limitations for the investigation into this classification of miRNA regulation.

9.2 Discovery of miRNA:miRNA Interactions

The first description of a miRNA:miRNA interaction was from Lai et al. (2004), who demonstrated that miRNA pairs were formed between miR-5 and miR-6, and between miR-9 and miR-79. These dyads demonstrated a stronger binding capacity compared to the complementary pairing between their respective miRNA guide and passenger strands (Lai et al. 2004). From this research, it was suggested that miRNAs may bind to each other in order to influence miRNA stability and mRNA targeting (Guo et al. 2012; Lai et al. 2004). Although this study was performed using sequence alignment and not confirmed in vitro, it established the concept that miRNAs may bind to and regulate miRNAs and other noncoding RNAs. Subsequent work has since established that miRNA:miRNA interactions do indeed occur in vitro and have broad impacts on cell homeostasis. A summary of the mechanisms behind miRNA:miRNA interactions and associated diseases is shown in Table 9.1.

Table 9.1 Description of each type of miRNA:miRNA interaction, their proposed mechanism, and associated diseases

miRNA:miRNA type	Proposed mechanism(s)	Identified disease(s)
Direct	• Targeting of pri-miRNAs by nuclear miRNAs (Forrest et al. 2010; Tang et al. 2012; Zisoulis et al. 2012). • Impede microprocessor cleavage (Wang et al. 2014, 2018a). • Complementary sequences in two mature miRNAs (Chen et al. 2011; Lai et al. 2004).	• Hepatocellular Carcinoma (Wang et al. 2018a). • Cardiac disease (Wang et al. 2014).
Indirect	• miRNAs control Transcription Factors, Promoters, and epigenetic controllers (Jia et al. 2016; Sylvestre et al. 2007; van Rooij et al. 2009). • miRNA control of the miRNA biogenesis components (Leonov et al. 2015; Wang et al. 2018b).	• Lung Cancer (Borzi et al. 2017). • Epithelial ovarian cancer (Wang et al. 2018b). • Colon cancer (Yu et al. 2015).
Global	• Culmination of changes as a result of promoter, transcription factor, and gene regulation (Matkovich et al. 2013). • Secondary regulatory pathways (Ooi et al. 2017).	• Ovarian cancer (Shahab et al. 2012). • Pulmonary hypertension (Bertero et al. 2014).

9.3 Direct miRNA:miRNA Interactions

Direct miRNA:miRNA regulation involves the binding and regulation of one miRNA by another, either in its mature form in the cytoplasm or pri-miRNA form in the nucleus. This section will discuss several examples of direct binding between two miRNAs, and how this relates to disease development.

9.3.1 Pri-miRNA:miRNA

Studies into miRNA regulation have found that selected pri-miRNAs contain sites for mature miRNAs, and that complementary binding between these two RNA forms results in a reduction in pri-miRNA processing and subsequent decrease in its associated miRNA. The first example of this showed that miR-424 and miR-503 target pri-miR-9 to control cell differentiation and lineage (Forrest et al. 2010). The main role of miR-9 is to maintain a non-differentiated cell state. However, miR-424 and miR-503 are in opposition to miR-9, as they are both pro-differentiative miRNAs. The downregulation of pri-miR-9 by miR-424 and miR-503, therefore, promotes cell lineage commitment and differentiation (Forrest et al. 2010). The targeting of pri-miR-9 by these two miRNAs implies that this interaction occurs in the nucleus, however, this aspect was not explored by the original authors. This example highlights how miRNA:miRNA interactions may participate in altering cell function and lineage.

Two major theories as to the actions of miRNA:miRNA interactions were discussed in a study by Tang et al. (2012). They demonstrated that miR-709 targeted pri-miR-15/16–1 in mice, and that the knockdown of Importin-8 (IPO8) prevented miR-709 from migrating into the nucleus to target pri-miR-15a/16–1. This implies that pri-miRNA targeting mature miRNAs are first produced in the cytoplasm and are then transported back into the nucleus to perform their regulatory role. The authors also established that miRNA:miRNA interactions have an influence on miRNA biogenesis.

miRNA:miRNA interactions have also been shown to have an autoregulatory aspect. In their study, Zisoulis et al. (2012) determined that in *C. elegans*, pri-let-7 was targeted by the mature form, let-7, to enhance its production, thus forming a positive feedback loop. Again, these results implied that mature miRNAs were both present and active in the nucleus, but also indicated that miRNAs may undergo autoregulation.

The studies mentioned observed that miRNAs were capable of binding to the primary miRNA form, but do not discuss or hypothesise on the mechanism by which miRNA binding impacts pri-miRNA expression. Two reports have since demonstrated that miRNA regulation is likely through the inhibition of Microprocessor attachment and processing. The first was performed in murine cardiomyocytes, where miR-361 targeted pri-miR-484 and prevented Drosha cleavage (Wang et al.

2014). In this case, a decrease in miR-361 directed targeting of pri-miR-484. Since miR-484 is influential in cardiomyocyte apoptosis, this pathway contributed to an anti-apoptotic state and was found to be associated with cardiac diseases, such as myocardial infarction (Wang et al. 2014).

The second study focused on hepatic cells, where under normal physiological conditions miR-122 targeted pri-miR-21 (Wang et al. 2018a). The authors directly demonstrated that due to the proximity of the miR-122 binding site to the Drosha cleavage junction, the interaction between these two miRNAs interrupted Drosha binding and resulted in the restricted expression of miR-21 in normal liver cells (Wang et al. 2018a). Due to the influential nature of miR-21, particularly in cancer, the maintenance of this relationship is essential to cellular homeostasis and preventing tumorigenesis. This is mostly observed through the loss or mutation in miR-122, which results in the decreased inhibition of miR-21. A higher level of miR-21 results in the targeted downregulation of Programmed Cell Death 4 (PDCD4), resulting in a loss of cell cycle control and promoting tumour development (Lu et al. 2008; Wang et al. 2018a). These studies demonstrate that mature miRNAs may impede Microprocessor by binding to a pri-miRNA in proximity to the Microprocessor cleavage site. Further investigations need to be conducted in order to determine whether this form of direct miRNA:miRNA regulation is universal across miRNAs, and its impact on disease.

9.3.2 Direct Binding between Mature miRNAs

Since the postulated binding of two miRNAs by Lai et al. (2004), there are few examples of this occurring in vitro. Chen et al. (2011) demonstrated that miR-107 and let-7 form a miRNA duplex, resulting in let-7 downregulation. The physical binding of these two miRNAs results in a miRNA complex that depends on the GAA internal loop structure of miR-107 (Chen et al. 2011). Since let-7 is a known tumour suppressor miRNA, its downregulation via this complex results in an increase in its oncogenic targets, and a subsequent increase in tumorigenic changes. However, this study brings forth the question of how two RNA-induced silencing complex (RISC)-bound mature miRNAs may recognise and bind to each other, and the implications of this on the RISC components. One study on miRNA cooperation suggested that amino acid residues within Argonaute (AGO) may interact to allow for two miRNAs to act together (Flamand et al. 2017). It may be that this mechanism, or similar, is in place to allow for the direct binding of two RISC-bound miRNAs. Another suggestion is that miRNA:miRNA interactions increase mature miRNA stability, and that this may be related to the observation that a miRNA is stabilised when bound to a canonical target (Park et al. 2017).

Since their discovery, direct miRNA:miRNA interactions have been a fascinating area of study. However, more work needs to be conducted to fully comprehend the scope and mechanism behind these interactions. For example, it is not yet fully known which miRNAs are transported into the nucleus or how this may occur. It is

also unknown whether miRNA binding to pri-miRNA is a widespread method of miRNA control. Additionally, the role of RISC and its capacity to bind to other miRNA-RISC complexes is not thoroughly researched.

9.4 Indirect miRNA:miRNA Interactions

miRNA:miRNA interactions may also occur via the indirect actions of another miRNA, adding another layer of complexity to cellular regulatory systems. This section will discuss the involvement of miRNAs in the several modes of indirect miRNA:miRNA interactions.

9.4.1 The Role of Transcriptional Regulation

One of the more explored mechanisms behind indirect miRNA:miRNA interactions is the miRNA-mediated control of transcriptional regulators, such as transcription factors and epigenetic markers. In this model, a miRNA targets the 3' UTR of a transcriptional regulator, thus altering its expression and the downstream levels of its targets, including other miRNAs (Song et al. 2015). It is expected that with further investigation, this form of miRNA regulation may be ubiquitously observed across cellular systems.

The first identified transcription factor-mediated miRNA:miRNA interaction was within murine cardiac muscle cells, whereby miR-208a regulated miR-208b and miR-499 (van Rooij et al. 2009). The slow myosin genes, Myosin Heavy Chain 7 (Myh7) and Myosin Heavy Chain 7b (Myh7b), contain the intronic miRNAs, miR-208b, and miR-499, whereas the fast myosin gene Myosin Heavy Chain 6 (Myh6) encodes for miR-208a. It was found that an increase in miR-208a suppressed the repressors of Myh7 and Myh7b, resulting in an increase in their transcription. This results in the subsequent production of miR-208b and miR-499, which suppress the repressors of the slow myosin genes. A positive feedback loop is then formed, as the slow myosin genes further activate miR-208b and miR-499 transcription. In this regulatory pathway, miR-208b is only upregulated by miR-208a in the presence of external stress stimuli, such as low thyroid hormone and high calcium levels (van Rooij et al. 2009). Thus, this miRNA:miRNA interaction allows for the accurate modulation of miRNA levels to alter physiological traits, in this case, muscle contraction.

Another example of a transcription factor-driven positive feedback loop is between the E2 Factor (E2F) family of transcription factors and miR-20a (Sylvestre et al. 2007). In this cycle, the miR-17 ~ 92a family, including miR-20a, targets the E2F genes. This is reciprocated by the E2F-driven activation of miR-20a via its promoter. Following this loop, an increase in miR-20a will lower E2F levels, resulting in a decrease in its activation. Thus, miR-20 is capable of modulating its

own expression via E2F. This allows for the adjustment of E2F levels to prevent apoptosis (Sylvestre et al. 2007). Also involved in this feedback network is the transcription factor and proto-oncogene MYC, as it forms a positive feedback loop with the E2F genes and transcriptionally regulates the miR-17 ~ 92a cluster (Aguda et al. 2008). Due to the role of this feedback system in maintaining cell cycle progression, its dysregulation results in increased proliferation and tumorigenesis (Pickering et al. 2009). The miR-20a/E2F/MYC feedback system has been demonstrated to impact the initiation and progression of glial tumours, with higher levels of miR-20a associated with malignancy stage (Gruszka et al. 2021). Similarly in prostate cancer, high levels of miR-20a were indicative of poor survival, and the presence of high risk, progressive disease (Stoen et al. 2021). Therefore, the feedback mechanism between E2F and miR-20a has far-reaching implications for cancer initiation and progression.

Transcriptional pathways involving miRNA:miRNA interactions have been shown to have implications on the control of the major oncogenes. In colorectal cancer, an oncogenic miRNA (oncomiR), miR-21, is involved in the regulation of miR-145 in order to amplify oncogenic changes (Yu et al. 2015). In this pathway, an increase in miR-21-induced K-Ras signalling increased the expression of Ras-responsive element-binding protein (RREBP), which in turn inhibited the expression of miR-145. Conversely, miR-145 negatively controlled miR-21 expression by targeting the miR-21 promoter, Activator Protein-1 (AP1), and reduced K-Ras signalling. miR-145 targets K-Ras and RREBP, its direct repressors, in order to increase its own production, thus forming a positive feedback loop. miR-21 also formed an indirect positive feedback loop with AP1. The loss of miR-145 via the miR-21 mediated pathway resulted in an increase in its target genes SRY-Box Transcription Factor 2 (SOX2), Nanog Homeobox (Nanog) and Octamer-Binding Protein 4 (Oct4), and lead to tumorigenic changes (Yu et al. 2015).

Another instance of this involves Mouse Double Minute 2 (MDM2) and p53 in lung cancer cells (Borzi et al. 2017). It was found that miR-660 repressed MDM2, which resulted in an increase in p53, a subsequent increase in miR-486-5p, and the miR-29 and miR-34 miRNA families (Borzi et al. 2017). This has implications on carcinogenesis, as the instability of p53 affects the phosphatidylinositol 3-kinase (PI3K)/protein kinase B (AKT) pathway, and other major cancer processes. It has been suggested that this axis could be targeted by therapeutics in order to stabilise p53 and restrict tumour growth (Borzi et al. 2017). Therefore, this pathway demonstrates the wider impact of miRNA regulation and how these may influence disease development through the involvement of key drivers of cancer.

Other indirect miRNA:miRNA interactions involve changes in methylation patterns due to the miRNA-directed targeting of DNA methyltransferases. In this case, miR-29b negatively regulates DNA methyltransferase 3 Beta (DMNT3B), which alters the methylation pattern of the promoter for miR-195. This results in an increase in miR-195 production (Jia et al. 2016). This pathway was significant in tongue squamous cell carcinoma, as both of these miRNAs have been shown to have tumour suppressive properties, and may be a suitable axis for targeted therapeutics (Jia et al. 2016).

9.4.2 The Role of the Biogenesis Components

Several miRNA:miRNA interactions have been found to occur through the regulation of the miRNA biogenesis components. Experimentally, very few miRNA sites within the 3' UTR's of the biogenesis components have been validated (Chou et al. 2018; Kishore et al. 2011). It is expected that the targeted regulation of a member of the biogenesis pathway, such as Dicer or AGO2, would result in a global decrease in miRNAs. This has yet to be shown, and current studies only show a limited number of miRNAs to be affected. However, key relationships have been uncovered that are associated with disease.

The first of these was observed in epithelial ovarian cancer, where the targeting of Dicer by miR-98-5p decreased miR-152 expression (Wang et al. 2018b). This has consequences on chemotherapy resistance, as high levels of miR-98-5p and consequently low levels of miR-152 results in the upregulation of the DNA repair gene RAD51 recombinase (RAD51) (Wang et al. 2018b). Testing of this pathway in mouse models concluded that the introduction of miR-152 and treatment with cisplatin resulted in greater decreased cell proliferation compared to either of these treatments individually (Wang et al. 2018b). This implies that the interaction between these two miRNAs, via Dicer, has a critical role in carcinogenesis, and may be utilised as a potential therapeutic pathway.

Another example related to disease is the impact of miR-132 on AGO2 expression, and subsequent decrease in miR-221 and increase in miR-146a levels (Leonov et al. 2015). These miRNAs have a role in angiogenesis, as miR-132 is pro-angiogenic, miR-221 is anti-angiogenic, and miR-146a is related to inflammation (Leonov et al. 2015). Therefore, an increase in miR-132 and subsequent decrease in miR-221 results in an increase in blood vessel formation and the suppression of anti-angiogenic pathways (Leonov et al. 2015).

Another manner in which the biogenesis components are involved in the interaction between two miRNAs is the recently discovered phenomenon of Microprocessor transfer. It was found that the Dicer-independent miRNA, miR-451, is reliant on miR-144 for its production. It was observed that once Microprocessor had cleaved miR-144, it was transferred to miR-451 for its production. Shorter, or less optimal miRNAs were found to be located in clusters, and were thus more likely to undergo Microprocessor transfer (Fang and Bartel 2020). This dependence was optimised by both the presence of miR-144 with a full-length stem, and the base pair width between the two miRNAs (Fang and Bartel 2020; Shang et al. 2020). Two similar studies saw that the absence of miR-144 abrogated miR-451 expression, and the extension of the miR-451 stem loop was adequate to induce independent microprocessor cleavage (Fang and Bartel 2020; Shang et al. 2020). Application of this process across the whole spectrum of miRNAs within the cell (miRNAome) gives some explanation as to how shorter strand miRNAs are produced, and has implications on the evolution of miRNA and small hairpin loops (Shang et al. 2020).

This form of miRNA processing dependence has implications for Dicer expression. Dicer contains a binding site for miR-144 within its 3' UTR, while miR-451 is

Dicer independent (Kretov et al. 2020). The production of miR-144 induces miR-451 cleavage by Microprocessor. An increased level of miR-144 induces the downregulation of Dicer, which is advantageous for miR-451, due to its Dicer independence. This results in an increase in AGO bound miR-451, and a decrease in other canonical miRNAs. This process is observed in red blood cells, where miR-451 is the dominant miRNA (Kretov et al. 2020).

9.5 Global miRNA:miRNA Interactions

miRNAs and their targets are part of a dynamic system. Small changes in the abundance of a subset of miRNAs may have a cascading effect on mRNAs and miRNA control. miRNA:miRNA interactions have system wide consequences, and thus it is important that we explore the impacts of miRNA aberrations on the cellular environment.

Several studies have been conducted to elucidate the network-wide impacts of miRNA changes. Shahab et al. (2012) observed the response of miRNAs and mRNAs with the overexpression of miR-7 in ovarian cancer cells, allowing for the identification of primary and secondary regulated miRNAs and genes. This brings forward the question of how a singular miRNA may impact miRNAs and mRNAs in both a direct and indirect manner. It was postulated that indirect changes may be the result of variations in promoter or transcription factor activity, the dysregulation of mRNA containing intronic miRNAs, and changes in the transcription of miRNAs from intergenic regions (Shahab et al. 2012).

A pivotal study by Matkovich et al. (2013) investigated the impact of miR-499 and miR-378 on miRNAs and mRNAs in murine cardiac cells. The transgenic overexpression of miR-499 and miR-378 resulted in 17 dysregulated miRNAs (11 up, 6 down) and 49 miRNAs (18 up, 31 down), respectively (Matkovich et al. 2013). Some of the changes in miRNA expression can be explained by regulatory cascades, as 13 of the dysregulated miRNAs were encoded within the gene targets of miR-499 or miR-378. Of the dysregulated mRNAs associated with miR-499 overexpression, only 7.8% (76) were targets of miR-499 itself, while 31% (298) of the targets were linked to the 11 upregulated miRNAs. The remaining dysregulated mRNAs (595) are likely the result of the observed secondary miRNA changes (Matkovich et al. 2013). It was found that miR-378 indirectly affects miR-99 expression by targeting MAF bZIP Transcription Factor (MAF) and Retinoic Acid Receptor-Related Orphan Receptor A (RORA), therefore indirectly altering the expression of 31 miR-99 targets (Matkovich et al. 2013). This study was critical in understanding the impact of miRNA expression on the global cell system.

Another study on the wider implication of miRNA:miRNA interactions investigated miR-130/miR-301 in the context of pulmonary hypertension (Bertero et al. 2014). The elevated presence of miR-130/301, as observed in hypoxic conditions, suppresses Peroxisome Proliferator-Activated Receptor Gamma (PPARγ), which in turn decreases Apelin, miR-322, miR-503, and increases Fibroblast Growth Factor

2 (FGF2) (Bertero et al. 2014). A decrease in PPARɣ also results in an increase in Signal Transducer And Activator Of Transcription 3 (STAT3), and a subsequent decrease in miR-204. Cumulatively, alterations in the mentioned miRNAs and genes promote a coordinated response to pulmonary distress and increases vascular remodelling. This aggregate effect is the result of miR-130/301 altering the expression of a wide range of miRNAs and genes, inducing a greater effect on cell functioning (Bertero et al. 2014).

9.5.1 The Wide Effect of a Small Set of miRNAs on Cell Functioning

There are several proposed mechanisms as to how a miRNA, or a miRNA family, may affect the spectrum of miRNAs and mRNAs within a cell system. One suggested theory is the presence of a "master regulator" miRNA, a miRNA that induces a change in an expanded network of genes and miRNAs, which results in a coordinated response to a stimulus (Bertero et al. 2014; Tang et al. 2012). This concept may well apply to many tissue types or cell systems, but is yet to be fully documented.

This concept was first proposed by Tang et al. (2012) in their investigation of miR-15a/16–1 regulation via miR-709. The authors introduced the idea of a miRNA hierarchy, whereby a group of miRNAs, or master regulators, conduct broader post-transcriptional miRNA control. The implication of this is the creation of a miRNA cascade, whereby the control of a miRNA by another has secondary and even tertiary effects on wider miRNA expression.

Another study demonstrated the actions of coordinated miRNA responses to drive cell processes toward a certain phenotype. In this model, miR-130/301 expression decreased PPARɣ expression, resulting in the repression of apelin, thus decreasing miR-424, miR-322, and miR-503. A decrease in PPARɣ expression also increased STAT3 levels, and ultimately decreased miR-204 expression (Bertero et al. 2014). By influencing these two pathways, miR-130/301 synchronise a response to pulmonary hypertension to increase vascular remodelling (Bertero et al. 2014).

This idea was also discussed in a study by Ooi et al. (2017) who studied the effects of knocked down levels of miR-34 on murine cardiac ventricles and cell lines with the aim to determine the primary and secondary changes in miRNAs in cardiac pathology. A degree of coordination was found between the expression of miRNAs and their interaction with transcription factors (Ooi et al. 2017). This indicates that secondary miRNA changes may be the result of alterations in transcription, and that these may act in a cooperative manner to amplify a response to a stimulus. Thus, it is evident that master regulator miRNAs have expanded networks and roles, and that these changes may be additive beyond the influence of just one miRNA.

9.6 miRNA:miRNA Dysregulation

There are several postulated theories as to how miRNAs are dysregulated within disease, including modifications to the miRNA sequence, changes in the biogenesis components, or the expression of regulatory factors.

Mutations within a miRNA sequence have a direct impact on its site-directed targeting of genes. Single nucleotide polymorphisms (SNPs) may occur within the seed region of the miRNA, which is responsible for target recognition and binding (Lewis et al. 2003). Additionally, isoforms of miRNAs, termed isomiRs, may also alter the sequence of the seed region via the addition of nucleotides at its 3' or 5′ end. IsomiRs and miRNA SNPs have also been found to be associated with disease development (Bofill-De Ros et al. 2020). It is not currently known whether miRNA:miRNA interactions occur via the seed region, or if isomiRs have a role in altering miRNA:miRNA regulation. However, mutations in this region alter miRNA expression and mRNA targeting, and thus have reverberating effects on mRNA regulation. This is therefore likely to impact miRNA:miRNA interactions (Króliczewski et al. 2018).

It was previously found that alterations in the IIIa or IIIb domains of Dicer impact strand selection in miRNA biogenesis (Vedanayagam et al. 2019). Mutations within these two domains enrich for 3p miRNAs, and alter a miRNA's 5p to 3p ratio. This has implications on gene targeting, as the 3p and 5p strands of a miRNA have a different and distinct set of targets (Vedanayagam et al. 2019). For example, in endometrial cancer, changes to the distribution of 3p and 5p miRNAs altered the let-7, miR-15/16, miR-29, miR-101, and miR-17 miRNA families, resulting in the loss of target gene repression (Vedanayagam et al. 2019). This phenomenon also brings forth the question as to how miRNA:miRNA interactions, either direct or indirect, are affected by changes in miRNA strand ratios.

Loss of function mutations in Exportin 5 (XPO5) have also been shown to impact mature miRNA expression. With XPO5 mutation, there is a decrease in pre-miRNA transport from the nucleus into the cytoplasm, resulting in a decrease in the expression of mature miRNAs (Kim et al. 2016). By altering overall miRNA levels, this mutation impacts miRNA-directed mRNA targeting and has reverberating consequences on cell function. It is also suggested that alterations in the biogenesis components, such as this, have ramifications on miRNA:miRNA interactions, whether direct or indirect. It is therefore important to consider the impact of mutations and alterations in the miRNA biogenesis machinery on the overall miRNA and mRNA landscape, and how this might contribute to cancer development (Hata and Kashima 2016).

It is also pivotal to consider the role of super enhancers in the exploration of miRNA:miRNA interactions. Super enhancers are genomic loci that contain several enhancer elements that respond to multiple transcription factors, and are generally responsible for the cell-specific expression of miRNAs and genes (Matsuyama and Suzuki 2019). Changes in the super enhancer region are responsible for both tumour suppressive and oncogenic changes, and thus are vital in cancer biology (Matsuyama

and Suzuki 2019). A broader, systems-level understanding of the miRNAs and mRNAs affected by super enhancers may uncover miRNA:miRNA pathways in both a homeostatic and a pathogenic-related context.

9.7 How Can miRNA:miRNA Interactions Be Utilised for Cancer Therapy?

Further investigation into miRNA:miRNA interactions provides another avenue for therapeutic design and development. The discovery of miRNA:miRNA interactions results in the identification of their direct and indirect targets, which can be integrated to create regulatory networks. This may then be used to predict the downstream effect of miRNA changes, or identify therapeutic targets and potential biomarkers (Cilek et al. 2017). This approach has been used by both Liu and Ye (2019) and Lapa et al. (2019) to incorporate mRNA, long non-coding RNA (lncRNA), and miRNA changes in laryngeal squamous cell carcinoma (LSCC) to identify hub genes or master regulator miRNAs. Additionally, miRNA:miRNA networks have also been used to identify miRNA changes as a result of common therapeutic treatments such as Trastuzumab in Breast Cancer (Cilek et al. 2017), Cisplatin in Ovarian Cancer (Wang et al. 2018b), or experimental therapies such as those against miR-34 in cardiac disease (Ooi et al. 2017). By investigating the wider cell context of miRNA:miRNA and mRNA:miRNA interactions, scientists are better able to identify off-target effects of newly designed therapeutics, especially those that target aberrant miRNA expression. More research is needed to unravel the complexity of miRNA:miRNA interactions and how they may be specific to cell type.

9.8 Current Limitations to miRNA:miRNA Discovery

The investigation into miRNA:miRNA interactions is in its infancy. There are limitations to current methods, both computational and biological, that need to be addressed in order for research into this miRNA regulatory mechanism to move forward. These include the incorporation of cell specificity into mRNA:miRNA and miRNA:miRNA interactions, and in vitro methods.

9.8.1 Cell Specificity

The miRNA and target relationships, for the majority, are exclusive to cell type, and thus this individuality is extended to miRNA:miRNA interactions (Salmanidis et al.

2014). Currently, common miRNA target algorithms do not consider cell specificity in their predictions (Rock et al. 2019). In addition, the presence of isomiRs and the distribution of miRNAs within the nucleus are also cell specific, and can alter which genes are targeted within a cell system (Nam et al. 2014). Therefore, information taken from miRNA:target databases, such as TargetScan (Agarwal et al. 2015) and miRTarBase (Huang et al. 2019) may not reflect the relationships present in a cell line or tissue of interest, and lead to inaccurate findings. Caution must be taken in creating miRNA:miRNA networks, as the pathways and connections found in one cell type cannot be directly applied to another. This issue of cell specificity is an ongoing area of research, both in the bioinformatic and biological sciences.

9.8.2 Identification

Current identification of miRNA:miRNA interactions has relied heavily on miRNA sequencing (miRSeq) and miRNA microarray methods. Microarray identification allows for the detection of a distinct set of miRNAs, congruent with current miRBase annotations. However, miRSeq techniques allow for the discovery of novel miRNAs, isomiRs, and sequence variants that may have a role in homeostatic and pathogenic miRNA:miRNA interactions (Grillone et al. 2020). It is recommended that miRSeq be paired with mRNASeq, as this allows for the establishment of miRNA:miRNA:mRNA networks based on physiological changes.

To detect changes in miRNAs, many studies have overexpressed a set miRNA using transfection or transgenic systems. However, this does not allow for the inference of biologically relevant changes in miRNAs in response to the introduced miRNAs. This is because the introduction of a miRNA into a system at high levels induces competition with the endogenous miRNAs for available AGO (Khan et al. 2009). This results in a decrease in endogenous miRNA expression and consequently, an increase in the levels of endogenous miRNA targets (Khan et al. 2009). It is therefore important to consider this phenomenon when uncovering miRNA:miRNA interactions, as changes in miRNA expression may be due to the transfection method, rather than biologically relevant miRNA:miRNA interactions.

Several examples of miRNA:miRNA interactions were found to be between a mature miRNA and a pri-miRNA within the nucleus (Forrest et al. 2010; Tang et al. 2012; Wang et al. 2014, 2018a; Zisoulis et al. 2012). However, this has not been described as a widespread phenomenon as this requires the annotation of pri-miRNAs, which are highly transient in nature (Kim et al. 2017). Researchers have previously used Drosha knock down techniques to identify its substrate pri-miRNA, which has ultimately identified 60% of the pri-miRNAs for conserved annotated miRNAs across humans and mice (Chang et al. 2015; Kim et al. 2017). Several studies have also attempted to identify the transcriptional start sites of known miRNAs, or used targeted primers 100 bp up and downstream of the precursor to determining the pri-miRNA strand (Conrad et al. 2020). These methods restrict the detection of regulatory sites, transcriptional or otherwise, that occur outside of the

defined region. A full annotation of pri-miRNAs would be required to elucidate the impact of miRNAs on pri-miRNA regulation, and uncover how ubiquitous this mode of miRNA:miRNA interaction is in cellular systems.

9.9 Conclusions

This chapter discussed several manners in which miRNAs may control the expression of one another, and the impact of this on downstream cell functions. Many identified miRNA:miRNA interactions occur between two specific miRNAs, or miRNA families, but the systems level role of these powerful regulators also needs to be considered. We need to re-evaluate the current canonical view of miRNAs to incorporate their role in wider miRNA and mRNA regulation. Several outstanding questions need to be addressed to fully appreciate the extent to which miRNAs control the miRNAome. This includes investigation into the extent of pri-miRNA targeting by nuclear miRNAs, the elucidation of cell-specific miRNA targets, and the incorporation of bioinformatic techniques to identify key pathways. Currently known miRNA:miRNA interactions were found to have a significant role in disease development, including cancer, and as such it is predicted that this form of regulation is more profound than previously thought. Awareness of the interplay between miRNAs is vital, not only to understanding molecular cell functioning, but also in the development of future therapeutics and biomarkers.

Acknowledgments MH is supported by an Australian Research Training Program Scholarship and a Translational Cancer Research Network PhD Scholarship Top-Up Award, supported by the Cancer Institute NSW.

References

Agarwal V, Bell GW, Nam J-W, Bartel DP (2015) Predicting effective microRNA target sites in mammalian mRNAs. eLife 4:e05005. https://doi.org/10.7554/eLife.05005

Aguda BD, Kim Y, Piper-Hunter MG, Friedman A, Marsh CB (2008) MicroRNA regulation of a cancer network: consequences of the feedback loops involving miR-17-92, E2F, and Myc. Proc Natl Acad Sci 105(50):19678–19683

Bertero T, Lu Y, Annis S, Hale A, Bhat B, Saggar R, Saggar R, Wallace WD, Ross DJ, Vargas SO, Graham BB, Kumar R, Black SM, Fratz S, Fineman JR, West JD, Haley KJ, Waxman AB, Chau BN, Cottrill KA, Chan SY (2014) Systems-level regulation of microRNA networks by miR-130/301 promotes pulmonary hypertension. J Clin Invest 124(8):3514–3528. https://doi.org/10.1172/JCI74773

Bofill-De Ros X, Yang A, Gu S (2020) IsomiRs: expanding the miRNA repression toolbox beyond the seed. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms 1863(4):194373

Borzi C, Calzolari L, Centonze G, Milione M, Sozzi G, Fortunato O (2017) Mir-660-p53-mir-486 network: a new key regulatory pathway in lung tumorigenesis. Int J Mol Sci 18(1). https://doi.org/10.3390/ijms18010222

Chang T-C, Pertea M, Lee S, Salzberg SL, Mendell JT (2015) Genome-wide annotation of microRNA primary transcript structures reveals novel regulatory mechanisms. Genome Res 25(9):1401–1409

Chen PS, Su JL, Cha ST, Tarn WY, Wang MY, Hsu HC, Lin MT, Chu CY, Hua KT, Chen CN, Kuo TC, Chang KJ, Hsiao M, Chang YW, Chen JS, Yang PC, Kuo ML (2011) miR-107 promotes tumor progression by targeting the let-7 microRNA in mice and humans. J Clin Invest 121(9): 3442–3455. https://doi.org/10.1172/JCI45390

Chou C-H, Shrestha S, Yang C-D, Chang N-W, Lin Y-L, Liao K-W, Huang W-C, Sun T-H, Tu S-J, Lee W-H (2018) miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. Nucleic Acids Res 46(D1):D296–D302

Cilek EE, Ozturk H, Gur Dedeoglu B (2017) Construction of miRNA-miRNA networks revealing the complexity of miRNA-mediated mechanisms in trastuzumab treated breast cancer cell lines. PLoS One 12(10):e0185558. https://doi.org/10.1371/journal.pone.0185558

Conrad T, Ntini E, Lang B, Cozzuto L, Andersen JB, Marquardt JU, Ponomarenko J, Tartaglia GG, Orom UAV (2020) Determination of primary microRNA processing in clinical samples by targeted pri-miR-sequencing. bioRxiv

Fang W, Bartel DP (2020) MicroRNA clustering assists processing of suboptimal MicroRNA hairpins through the action of the ERH protein. Mol Cell 78(2):289–302.e286. https://doi.org/10.1016/j.molcel.2020.01.026

Flamand MN, Gan HH, Mayya VK, Gunsalus KC, Duchaine TF (2017) A non-canonical site reveals the cooperative mechanisms of microRNA-mediated silencing. Nucleic Acids Res. https://doi.org/10.1093/nar/gkx340

Forrest AR, Kanamori-Katayama M, Tomaru Y, Lassmann T, Ninomiya N, Takahashi Y, de Hoon MJ, Kubosaki A, Kaiho A, Suzuki M, Yasuda J, Kawai J, Hayashizaki Y, Hume DA, Suzuki H (2010) Induction of microRNAs, mir-155, mir-222, mir-424 and mir-503, promotes monocytic differentiation through combinatorial regulation. Leukemia 24(2):460–466. https://doi.org/10.1038/leu.2009.246

Grillone K, Riillo C, Scionti F, Rocca R, Tradigo G, Guzzi PH, Alcaro S, Di Martino MT, Tagliaferri P, Tassone P (2020) Non-coding RNAs in cancer: platforms and strategies for investigating the genomic "dark matter". J Exp Clin Cancer Res 39(1):1–19

Gruszka R, Zakrzewski K, Liberski PP, Zakrzewska M (2021) mRNA and miRNA expression analyses of the MYC/E2F/miR-17-92 network in the Most common pediatric brain tumors. Int J Mol Sci 22(2):543

Guo L, Sun B, Wu Q, Yang S, Chen F (2012) miRNA-miRNA interaction implicates for potential mutual regulatory pattern. Gene 511(2):187–194. https://doi.org/10.1016/j.gene.2012.09.066

Hata A, Kashima R (2016) Dysregulation of microRNA biogenesis machinery in cancer. Crit Rev Biochem Mol Biol 51(3):121–134. https://doi.org/10.3109/10409238.2015.1117054

Huang H-Y, Lin Y-C-D, Li J, Huang K-Y, Shrestha S, Hong H-C, Tang Y, Chen Y-G, Jin C-N, Yu Y, Xu J-T, Li Y-M, Cai X-X, Zhou Z-Y, Chen X-H, Pei Y-Y, Hu L, Su J-J, Cui S-D, Wang F, Xie Y-Y, Ding S-Y, Luo M-F, Chou C-H, Chang N-W, Chen K-W, Cheng Y-H, Wan X-H, Hsu W-L, Lee T-Y, Wei F-X, Huang H-D (2019) miRTarBase 2020: updates to the experimentally validated microRNA–target interaction database. Nucleic Acids Res 48(D1): D148–D154. https://doi.org/10.1093/nar/gkz896

Jia L-F, Zheng Y-F, Lyu M-Y, Huang Y-P, Gan Y-H (2016) miR-29b upregulates miR-195 by targeting DNMT3B in tongue squamous cell carcinoma. Science Bulletin 61(3):212–219. https://doi.org/10.1007/s11434-016-1001-6

Khan AA, Betel D, Miller ML, Sander C, Leslie CS, Marks DS (2009) Transfection of small RNAs globally perturbs gene regulation by endogenous microRNAs. Nat Biotechnol 27(6):549–555

Kim B, Jeong K, Kim VN (2017) Genome-wide mapping of DROSHA cleavage sites on primary microRNAs and noncanonical substrates. Mol Cell 66(2):258–269. e255

Kim YK, Kim B, Kim VN (2016) Re-evaluation of the roles of DROSHA, export in 5, and DICER in microRNA biogenesis. Proc Natl Acad Sci U S A 113(13):E1881–E1889. https://doi.org/10.1073/pnas.1602532113

Kishore S, Jaskiewicz L, Burger L, Hausser J, Khorshid M, Zavolan M (2011) A quantitative analysis of CLIP methods for identifying binding sites of RNA-binding proteins. Nat Methods 8(7):559–564

Kretov DA, Walawalkar IA, Mora-Martin A, Shafik AM, Moxon S, Cifuentes D (2020) Ago2-dependent processing allows miR-451 to evade the global MicroRNA turnover elicited during erythropoiesis. Mol Cell 78(2):317–328.e316. https://doi.org/10.1016/j.molcel.2020.02.020

Króliczewski J, Sobolewska A, Lejnowski D, Collawn JF, Bartoszewski R (2018) microRNA single polynucleotide polymorphism influences on microRNA biogenesis and mRNA target specificity. Gene 640:66–72

Lai EC, Wiel C, Rubin GM (2004) Complementary miRNA pairs suggest a regulatory role for miRNA:miRNA duplexes. RNA 10(2):171–175

Lapa RML, Barros-Filho MC, Marchi FA, Domingues MAC, de Carvalho GB, Drigo SA, Kowalski LP, Rogatto SR (2019) Integrated miRNA and mRNA expression analysis uncovers drug targets in laryngeal squamous cell carcinoma patients. Oral Oncol 93:76–84

Leonov G, Shah K, Yee D, Timmis J, Sharp TV, Lagos D (2015) Suppression of AGO2 by miR-132 as a determinant of miRNA-mediated silencing in human primary endothelial cells. Int J Biochem Cell Biol 69:75–84. https://doi.org/10.1016/j.biocel.2015.10.006

Lewis BP, Shih I-h, Jones-Rhoades MW, Bartel DP, Burge CB (2003) Prediction of mammalian microRNA targets. Cell 115(7):787–798

Liu Y, Ye F (2019) Construction and integrated analysis of crosstalking ceRNAs networks in laryngeal squamous cell carcinoma. PeerJ 7:e7380

Lu Z, Liu M, Stribinskis V, Klinge CM, Ramos KS, Colburn NH, Li Y (2008) MicroRNA-21 promotes cell transformation by targeting the programmed cell death 4 gene. Oncogene 27(31): 4373–4379. https://doi.org/10.1038/onc.2008.72

Matkovich SJ, Hu Y, Dorn GW 2nd (2013) Regulation of cardiac microRNAs by cardiac microRNAs. Circ Res 113(1):62–71. https://doi.org/10.1161/CIRCRESAHA.113.300975

Matsuyama H, Suzuki HI (2019) Systems and synthetic microRNA biology: from biogenesis to disease pathogenesis. Int J Mol Sci 21(1):132. https://doi.org/10.3390/ijms21010132

Nam J-W, Rissland OS, Koppstein D, Abreu-Goodger C, Jan CH, Agarwal V, Yildirim MA, Rodriguez A, Bartel DP (2014) Global analyses of the effect of different cellular contexts on microRNA targeting. Mol Cell 53(6):1031–1043

Ooi JYY, Bernardo BC, Singla S, Patterson NL, Lin RCY, McMullen JR (2017) Identification of miR-34 regulatory networks in settings of disease and antimiR-therapy: implications for treating cardiac pathology and other diseases. RNA Biol 14(5):500–513. https://doi.org/10.1080/15476286.2016.1181251

Park JH, Shin SY, Shin C (2017) Non-canonical targets destabilize microRNAs in human Argonautes. Nucleic Acids Res 45(4):1569–1583. https://doi.org/10.1093/nar/gkx029

Pickering MT, Stadler BM, Kowalik TF (2009) miR-17 and miR-20a temper an E2F1-induced G1 checkpoint to regulate cell cycle progression. Oncogene 28(1):140–145

Rock LD, Minatel BC, Marshall EA, Guisier F, Sage AP, Barros-Filho MC, Stewart GL, Garnis C, Lam WL (2019) Expanding the transcriptome of head and neck squamous cell carcinoma through novel MicroRNA discovery. Front Oncol 9:1305

Salmanidis M, Pillman K, Goodall G, Bracken C (2014) Direct transcriptional regulation by nuclear microRNAs. Int J Biochem Cell Biol 54:304–311

Shahab SW, Matyunina LV, Mittal VK, Wang L, Hill CG, Walker LD, McDonald JF (2012) MicroRNAs indirectly regulate other microRNAs in ovarian cancer cells. British J Med Med Res 2(2):172

Shang R, Baek SC, Kim K, Kim B, Kim VN, Lai EC (2020) Genomic clustering facilitates nuclear processing of suboptimal Pri-miRNA loci. Mol Cell 78(2):303–316.e304. https://doi.org/10.1016/j.molcel.2020.02.009

Song C, Xu Z, Jin Y, Zhu M, Wang K, Wang N (2015) The network of microRNAs, transcription factors, target genes and host genes in human renal cell carcinoma. Oncol Lett 9(1):498–506. https://doi.org/10.3892/ol.2014.2683

Stoen MJ, Andersen S, Rakaee M, Pedersen MI, Ingebriktsen LM, Donnem T, Lombardi AP, Kilvaer TK, Busund L-TR, Richardsen E (2021) Overexpression of miR-20a-5p in tumor epithelium is an independent negative prognostic indicator in prostate cancer–a multi-institutional study. Cancers 13(16):4096

Sylvestre Y, De Guire V, Querido E, Mukhopadhyay UK, Bourdeau V, Major F, Ferbeyre G, Chartrand P (2007) An E2F/miR-20a autoregulatory feedback loop. J Biol Chem 282(4): 2135–2143. https://doi.org/10.1074/jbc.M608939200

Tang R, Li L, Zhu D, Hou D, Cao T, Gu H, Zhang J, Chen J, Zhang CY, Zen K (2012) Mouse miRNA-709 directly regulates miRNA-15a/16-1 biogenesis at the posttranscriptional level in the nucleus: evidence for a microRNA hierarchy system. Cell Res 22(3):504–515. https://doi.org/10.1038/cr.2011.137

van Rooij E, Quiat D, Johnson BA, Sutherland LB, Qi X, Richardson JA, Kelm RJ Jr, Olson EN (2009) A family of microRNAs encoded by myosin genes governs myosin expression and muscle performance. Dev Cell 17(5):662–673. https://doi.org/10.1016/j.devcel.2009.10.013

Vedanayagam J, Chatila WK, Aksoy BA, Majumdar S, Skanderup AJ, Demir E, Schultz N, Sander C, Lai EC (2019) Cancer-associated mutations in DICER1 RNase IIIa and IIIb domains exert similar effects on miRNA biogenesis. Nat Commun 10(1):3682. https://doi.org/10.1038/s41467-019-11610-1

Wang D, Sun X, Wei Y, Liang H, Yuan M, Jin F, Chen X, Liu Y, Zhang CY, Li L, Zen K (2018a) Nuclear miR-122 directly regulates the biogenesis of cell survival oncomiR miR-21 at the posttranscriptional level. Nucleic Acids Res 46(4):2012–2029. https://doi.org/10.1093/nar/gkx1254

Wang K, Sun T, Li N, Wang Y, Wang JX, Zhou LY, Long B, Liu CY, Liu F, Li PF (2014) MDRL lncRNA regulates the processing of miR-484 primary transcript by targeting miR-361. PLoS Genet 10(7):e1004467. https://doi.org/10.1371/journal.pgen.1004467

Wang Y, Bao W, Liu Y, Wang S, Xu S, Li X, Li Y, Wu S (2018b) miR-98-5p contributes to cisplatin resistance in epithelial ovarian cancer by suppressing miR-152 biogenesis via targeting Dicer1. Cell Death Dis 9(5):447. https://doi.org/10.1038/s41419-018-0390-7

Yu Y, Nangia-Makker P, Farhana L, S GR, Levi E, Majumdar AP (2015) miR-21 and miR-145 cooperation in regulation of colon cancer stem cells. Mol Cancer 14:98. https://doi.org/10.1186/s12943-015-0372-7

Zisoulis DG, Kai ZS, Chang RK, Pasquinelli AE (2012) Autoregulation of microRNA biogenesis by let-7 and Argonaute. Nature 486(7404):541–544. https://doi.org/10.1038/nature11134

Chapter 10
ClustMMRA v2: A Scalable Computational Pipeline for the Identification of MicroRNA Clusters Acting Cooperatively on Tumor Molecular Subgroups

Céline Hernandez, Gabriele Cancila, Olivier Ayrault, Andrei Zinovyev, and Loredana Martignetti

Abstract In recent cancer genomics programs, large-scale profiling of microRNAs has been routinely used in order to better understand the role of microRNAs in gene regulation and disease. To support the analysis of such amount of data, scalability of bioinformatics pipelines is increasingly important to handle larger datasets.

Here, we describe a scalable implementation of the clustered miRNA Master Regulator Analysis (clustMMRA) pipeline, developed to search for genomic clusters of microRNAs potentially driving cancer molecular subtyping. Genomically clustered microRNAs can be simultaneously expressed to work in a combined manner and jointly regulate cell phenotypes. However, the majority of computational approaches for the identification of microRNA master regulators are typically designed to detect the regulatory effect of a single microRNA.

We have applied the clustMMRA pipeline to multiple pediatric tumor datasets, up to a hundred samples in size, demonstrating very satisfying performances of the software on large datasets. Results have highlighted genomic clusters of microRNAs potentially involved in several subgroups of the different pediatric cancers or

C. Hernandez
Université Paris-Saclay, CEA, CNRS, Institute for Integrative Biology of the Cell (I2BC), Gif-sur-Yvette, France

G. Cancila · O. Ayrault
Institut Curie, PSL Research University, Paris, France

CNRS UMR 3347, INSERM U1021, Université Paris Sud, Université Paris-Saclay, Orsay, France

A. Zinovyev · L. Martignetti (✉)
Institut Curie, PSL Research University, Paris, France

INSERM U900, Paris, France

MINES ParisTech, PSL Research University, CBIO-Centre for Computational Biology, Paris, France
e-mail: Loredana.Martignetti@curie.fr

U. Schmitz et al. (eds.), *Systems Biology of MicroRNAs in Cancer*, Advances in Experimental Medicine and Biology 1385,
https://doi.org/10.1007/978-3-031-08356-3_10

specifically involved in the phenotype of a subgroup. In particular, we confirmed the cluster of microRNAs at the 14q32 locus to be involved in multiple pediatric cancers, showing its specific downregulation in tumor subgroups with aggressive phenotype.

Keywords MicroRNA clusters · Combinatorial targeting · Posttranscriptional regulation · Pediatric tumors · Molecular tumor subgroups

10.1 Introduction

MicroRNAs (miRNAs) are short regulatory RNAs discovered as important post-transcriptional repressors of gene expression in diverse biological contexts (Bartel 2004). These ~22 nt long molecules recognize and bind to partially complementary sites in the 3′ UTR of target transcripts. Even though their precise mechanism of action is still under investigation, the current idea is that miRNAs are able to negatively affect the protein output through a combination of mRNA destabilization or translational repression (Dexheimer and Cochella 2020). In plants, usually miRNAs have perfect or near-perfect antisense complementarity to their mRNA target, whereas in animals the complementarity is restricted to the 5′ region of the miRNA, in particular requiring a *seed* match of six nucleotides, around nucleotides two–seven.

Both computational (Friedman et al. 2009) and experimental studies (Lim et al. 2005; Baek et al. 2008; Selbach et al. 2008; Hendrickson et al. 2009; Liu and Wang 2019) have revealed that most human protein-coding genes are regulated by one or more miRNAs. Each miRNA has hundreds of target genes and a single gene can be targeted by several miRNAs. This suggests the existence of a combinatorial post-transcriptional regulatory layer which contributes to the complexity of gene expression patterns.

In humans and other mammals, miRNAs have been shown to be involved in the regulation of all essential cell functions from differentiation and proliferation to apoptosis (Bartel 2018) and to be aberrantly expressed in many diseases, including cancer (Calin and Croce 2006). Multiple studies have demonstrated the importance of miRNAs in all the cancer hallmarks defined by Hanahan and Weinberg (Van Roosbroeck and Calin 2017) and indicated that they might function as oncogenes or tumor suppressors (Cimmino et al. 2005; Ma et al. 2007; Valastyan et al. 2009). Further experimental evidence suggested that specific miRNAs may also have a role beyond the cancer onset and directly participate in cancer invasiveness and metastasis (Ma et al. 2007; Brabletz and Brabletz 2010). Indeed, miRNA profiles can distinguish not only between normal and cancerous tissue but they can also successfully classify different subgroups of a particular cancer (Yang et al. 2013; Rooj et al. 2017).

Identifying key miRNAs contributing to the genesis and development of a particular disease is the goal of many studies. However, most studies have focused their attention on single miRNA–target gene interactions, thereby neglecting the

combinatorial nature of gene regulation in higher eukaryotes. A recent work coupling bioinformatics and functional validations (Cursons et al. 2018) has precisely described a combinatorial regulatory circuit composed of a set of cooperating miRNAs controlling epithelial-to-mesenchymal transition (EMT).

With the aim of identifying regulation by multiple miRNA species acting cooperatively on pathological phenotypes, we recently designed the computational pipeline named Clustered MiRNA Master Regulator Analysis (ClustMMRA) (Cantini et al. 2019) to search for genomically clustered miRNAs potentially driving cancer subtyping, a comparison characterized by much lower variation than cancer versus normal conditions. Approximately 25% of human miRNA genes are organized in polycistronic genomic units or clusters (Altuvia et al. 2005), that contain two or more miRNA genes transcribed from physically adjacent sequences. Genomically clustered miRNAs can be simultaneously expressed and work in a combined manner, jointly regulating targets in close proximity of the protein–protein interaction network (Liang and Li 2007), or belonging to the same pathway (Tsang et al. 2010).

Among the best-studied miRNA clusters, the miR-17/92 cluster has been shown to promote tumorigenesis in multiple cancers, including B-cell lymphoma (He et al. 2005; Mu et al. 2009; Olive et al. 2009), small-cell lung cancer, colon cancer, neuroblastoma (NB), and medulloblastoma (MB) (Hayashita et al. 2005; Mestdagh et al. 2010; Uziel et al. 2009; Murphy et al. 2013). The emerging picture suggests that miR-17/92 plays a widespread role in tumorigenesis, but the specific miRNAs involved and the key targets regulated appear highly context dependent.

Other studies have shown that the cluster of miRNAs at the 14q32 locus is involved in multiple pediatric cancers, including desmoplastic MB (Lucon et al. 2013b), and NB (Soriano et al. 2019; Bhavsar et al. 2018). It was also shown to define clinically relevant molecular subgroups in osteosarcoma (Hill et al. 2017), glioblastoma (Shahar et al. 2016), and to be metastasis-suppressive in preclinical models of lung and liver metastasis (Oshima et al. 2019). Another well-studied cluster, the miR-183/96/182, has been shown to regulate cell survival, proliferation, and migration in MB (Weeraratne et al. 2012).

The clustMMRA pipeline allows for the unbiased identification of clusters of miRNAs that potentially drive cancer subtyping. In a previous study on breast cancer (Cantini et al. 2019), the results have highlighted key miRNA clusters driving the phenotype of different subgroups and the identification of miR-199/miR-214 as a novel cluster promoting the triple-negative breast cancer (TNBC) phenotype through its control of proliferation and EMT.

Here, we describe an improved implementation of the clustMMRA pipeline that achieves an important improvement in computational performance in both time and memory usage. This made it possible to use the pipeline to efficiently analyze datasets containing hundreds of samples.

We applied the clustMMRA pipeline to multiple pediatric cancers, including NB, MB and ependymoma (EPN), all diseases formed by the degeneration of primitive tissue during the embryonic development that present a high heterogeneity in terms of subgroups. The aim of our analysis is to highlight genomic clusters of miRNAs

potentially involved in several subgroups of the different diseases or specifically involved in the phenotype of a subgroup.

10.2 The clustMMRA Pipeline

The clustMMRA pipeline is a multi-step workflow that requires input miRNA/mRNA expression profiles from matched tumor samples classified in different subgroups according to subgroup-specific gene signatures. The final output of clustMMRA provides key miRNA clusters contributing to the regulation of particular subgroups of the disease.

The workflow of clustMMRA (see Fig. 10.1) consists of subsequent steps: (i) differential expression analysis of clustered miRNAs; (ii) target enrichment analysis; and (iii) network analysis.

In step (i), the subgroup-specific expression of each miRNA is assessed by Kolmogorov–Smirnov (KS) statistical test and fold change cutoff. While a miRNA cluster is usually transcribed as a single unit, the expression of mature miRNAs in the same cluster might not be highly correlated due to regulatory events in the maturation processes. The genomic locations of miRNAs were retrieved from

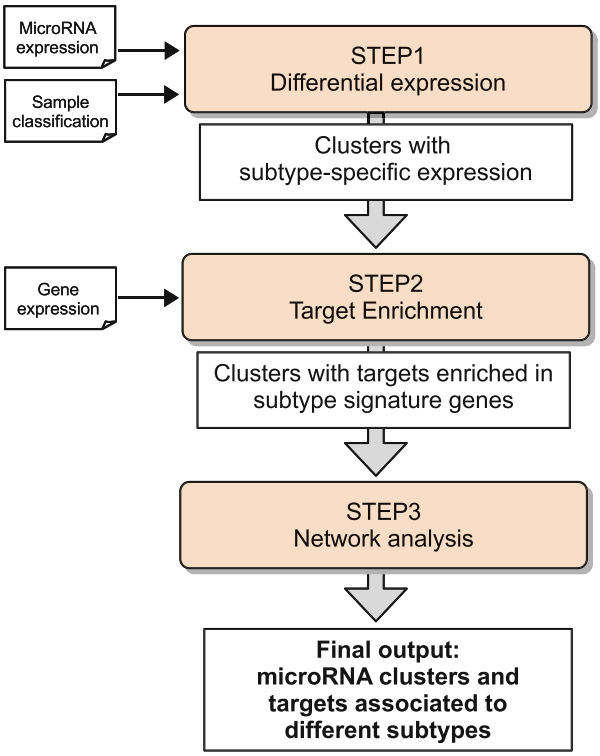


Fig.10.1 Schematic representation of the Clustered microRNA Master Regulator Analysis (ClustMMRA) workflow. The diagram shows the data required as inputs, the three analytical steps with their respective outputs and the final output of the pipeline

the miRBase v18 database (Kozomara et al. 2019). As in previous studies (Altuvia et al. 2005; Marco et al. 2013), co-clustered miRNAs were defined as miRNA genes located within 10 kb of distance in the same chromosome and on the same strand. Clusters having at least two miRNAs with subgroup-specific expression change in the same direction (both upregulated or downregulated) are selected for step (ii).

In step (ii), miRNA clusters that have their predicted targets enriched for the gene signature of the corresponding subgroup are selected. Individual miRNA targets are predicted using four different databases, namely TargetScan (Agarwal et al. 2015), DoRiNA (Anders et al. 2012), PicTar (Krek et al. 2005), and PITA (Kertesz et al. 2007), plus an experimental one miRTarBase (Huang et al. 2020) and requesting prediction from at least two databases.

The set of targets of a cluster has been defined as the union of the targets of individual miRNAs.

To control for false positives, we tested the significance of the overlap between the targets of the cluster and the genes present in the signature based on a null model. The null model was generated taking 1000 random sets of genes of the same size as the target set and the true value of the intersection is tested, setting the 95th percentile of the null distribution as the threshold. The goal of steps (i) and (ii) is to identify co-clustered and co-expressed miRNAs potentially regulating a gene expression signature jointly, without necessarily having a high overlap in terms of target genes (Hausser and Zavolan 2014).

Finally, in step (iii) a miRNA–mRNA interaction network is constructed for each selected cluster using the ARACNE algorithm (Margolin et al. 2006; He et al. 2017) and including all the expressed genes. At this stage, we identify co-clustered miRNA modules and interacting genes, including indirect interactions, which are believed to participate in the phenotype of a given cancer subgroup (we call these modules *regulons*). Permitting indirect interactions can introduce spurious links between miRNAs and genes in the regulon. A Fisher's exact test was performed to evaluate the statistical significance of the overlap between the genes included in each regulon and the gene signature of the associated subgroup.

The implementation was reviewed in order to better document each computation step, and the global flow of information optimized, thus clarifying the implemented analytic process. Moreover, the version of third-party software used by the pipeline was updated, especially ARACNE (Margolin et al. 2006; He et al. 2017). This update removed the dependency on a supporting computing cluster. ClustMMRA can now run independently on a laptop with a recent processor and 32 Gb RAM.

10.3 ClustMMRA Usage

The primary input for clustMMRA is two tab-separated matrices containing expression values of genes and miRNAs in matched samples, with columns sorted according to the cancer subtype conditions.

The two normalized expression matrices have to be composed of the same samples and include sample IDs in the first column and gene or miRNA IDs in the first row. Samples can be classified in different subtypes based on external information or according to transcriptomic profiles. A typical unsupervised classification can be performed using hierarchical clustering on the most variant genes. To obtain a robust classification, a consensus clustering approach such as the Monti consensus algorithm can be applied (Monti et al. 2003; John et al. 2020). The miRNA expression matrix will be used in step (i) for selecting clusters differentially expressed in the different subtypes. Both miRNA and gene expression matrices will be used in step (iii) to build a miRNA–mRNA interaction network for each selected cluster based on mutual information between miRNA and gene expression.

The second required input consists of transcriptomic signatures for each cancer subtype. These are provided as single-column text files reporting the list of those genes that represent the considered subtype. In step (ii), these signatures are tested for enrichment of miRNA targets predicted by different databases.

The clustMMRA pipeline is implemented as a main command line program called run_*clustMMRA.sh*, which can be run on multiple computing platforms. The main program asks for parameters interactively and redirects the output of each step to the input of the next one. Intermediate output files are produced for each step of the pipeline. It invokes R functions for computing the first two steps of the pipeline and the third party ARACNE-AP software (Lachmann et al. 2016) in step (iii). The ARACNE algorithm represents one of the most effective tools to accomplish the goal of network reverse engineering. ARACNE-AP operates on a bootstrapped version of the input matrix, generating N mutual information networks for N bootstraps. It implements a consolidation step in which a consensus network is computed by estimating the statistical significance of the number of times a specific edge is detected across all bootstrap runs, based on a Poisson distribution and a Bonferroni correction on the obtained p-values.

The final output of clustMMRA consists of a network file (*regulon*) for each cluster associated to a given subtype. A network file is a tab delimited text file with the following four columns: miRNA ID, target ID, mutual information, and the corresponding p-value. This network includes co-clustered miRNAs and interacting genes connected by both direct and indirect interactions, which are believed to participate in the phenotype of a given cancer subtype.

10.4 Application to Pediatric Cancers

Here, we present the application of clustMMRA to miRNA and mRNA expression data of three embryonic tumors for which we collected miRNA/mRNA expression data from matched samples. ClustMMRA was applied to identify polycistronic miRNAs potentially involved in the regulation of cancer molecular subgroups of these tumors.

The first dataset (from GSE121513 and GSE85047) consists of gene and miRNA expression profiles of 95 NB samples classified in two patient subgroups based on *MYCN* gene status (not amplified vs. amplified). Even though *MYCN* amplification is a key indicator of poor prognosis, mechanisms by which *MYCN* promotes NB tumorigenesis are not fully understood. Given the clinical heterogeneity of NB, clusters of miRNAs associated to *MYCN* regulation may be of direct biological relevance for molecular subgroups.

The second dataset includes miRNA and mRNA expression from 26 patients of MB, a malignant pediatric brain tumor in childhood most commonly formed in the cerebellum. The current consensus recognizes four main MB subgroups (WNT, SHH, G3, and G4). While the WNT and SHH subgroups are characterized by clearly defined aberrant oncogenic activation of developmental pathways, genomic and transcriptomic approaches have failed to identify molecular aspects in G3 and G4. A recent proteomic study in MB (Forget et al. 2018) revealed a relatively low degree of concordance between the mRNA and protein expression in G3 and G4, suggesting a possibly important role in post-transcriptional regulation.

The third dataset (GSE21687) consists of 58 patient samples of EPN, a neural tumor that arises throughout the central nervous system (CNS). According to genomic, transcriptomic, and miRNA profiles (Johnson et al. 2010), EPN tumors are classified into nine molecular subgroups (A to I). The observation that different molecular profiles are able to classify EPNs into similar subgroups supports the notion that these subgroups are true biological entities.

We applied the clustMMRA pipeline to the three datasets separately. The analysis outputs those miRNA clusters potentially contributing to the regulation of molecular subgroups. The regulons associated by clustMMRA to a given subgroup contain multiple miRNAs of the genomic cluster and their links to targets in that subgroup.

The clustMMRA output shows that 44 of the 133 clusters included in the analysis are significantly associated with at least one disease subgroup, 6 in NB, 23 in MB, and 15 in EPN, respectively (Table 10.1). Of these, six clusters are involved in the subgrouping of three embryonic cancers, four clusters are involved in two of the three cancers and 18 clusters are involved only in one cancer. To understand which clusters are involved in several subgroups and which subgroups share the regulation by the same clusters, we have represented these results in a bipartite graph where two sets of nodes (miRNA clusters and cancer subgroups) are connected by links corresponding to the potential regulation predicted by clustMMRA (Fig. 10.2).

The network shows that two clusters, the C19MC cluster on chr19q13 and the C14CM cluster on chromosome 14, are involved in the regulation of numerous subgroups of the three studied cancers. The C19MC cluster is significantly associated with the regulation of the subgroup G4 in MB, the subgroups A, B, F, I in EPN, and the MYCN amplified and not amplified subgroups in NB. It is under-expressed in subgroups A, B, and I of EPN while it is over-expressed in subgroup F of EPN and subgroup G4 of MB. However, its expression is ambivalent in the NB subgroups, since some miRNAs of the cluster are over-expressed in MYCN amplified and not amplified subgroups and some others are under-expressed. The C14MC cluster is predicted to regulate the G3 and G4 subgroups in MB, the subgroups B, E, G, and I

Table 10.1 Clusters of miRNAs associated to pediatric cancer subgroups

Cluster ID	Expr. class	Cluster sign	Cluster members	Size
cl1189_chrX	EPN_A	up	miR-188-5p/miR-188-3p/miR-362-5p/miR-362-3p/miR-500a-5p/miR-500a-3p/miR-501-5p/miR-501-3p/miR-502-5p/miR-502-3p/miR-532-5p/miR-532-3p/miR-660-5p/miR-660-3p/miR-500b	15
cl1226_chrX	EPN_A	down	miR-513a-5p/miR-513a-3p/miR-506-3p/miR-506-5p/miR-507/miR-508-5p/miR-508-3p	7
cl228_chr11	EPN_A	down	miR-34b-5p/miR-34b-3p/miR-34c-5p/miR-34c-3p	4
cl491_chr17	EPN_A	down	miR-142-5p/miR-142-3p/miR-4736	3
cl590_chr19	EPN_A	down	miR-512-5p/miR-512-3p/miR-512-5p/miR-512-3p/miR-498/miR-520e/miR-515-5p/miR-515-3p/miR-519e-5p/miR-519e-3p/miR-520f/miR-515-5p/miR-515-3p/miR-519c-5p/miR-519c-3p/miR-520a-5p/miR-520a-3p/miR-526b-5p/miR-526b-3p/miR-519b-5p/miR-519b-3p/miR-525-5p/miR-525-3p/miR-523-5p/miR-523-3p/miR-518f-5p/miR-518f-3p/miR-520b/miR-518b/miR-526a/miR-520c-5p/miR-520c-3p/miR-518c-5p/miR-518c-3p/miR-524-5p/miR-524-3p/miR-517-5p/miR-517a-3p/miR-519d/miR-521/miR-520d-5p/miR-520d-3p/miR-517b-3p/miR-517-5p/miR-520 g/miR-516b-5p/miR-516b-3p/miR-526a/miR-518e-5p/miR-518e-3p/miR-518a-5p/miR-518a-3p/miR-518d-5p/miR-518d-3p/miR-516b-5p/miR-516b-3p/miR-518a-5p/miR-518a-3p/miR-517-5p/miR-517c-3p/miR-520 h/miR-521/miR-522-5p/miR-522-3p/miR-519a-5p/miR-519a-3p/miR-527/miR-516a-5p/miR-516a-3p/miR-516a-5p/miR-516a-3p/miR-519a-3p/miR-1323/miR-1283/miR-1283	75
cl611_chr2	EPN_A	down	miR-216a/miR-217	2
cl742_chr22	EPN_A	up	miR-130b-5p/miR-130b-3p/miR-301b	3
cl1161_chr9	EPN_B	up	miR-199b-5p/miR-199b-3p/miR-3154	3
cl1217_chrX	EPN_B	up	miR-424-5p/miR-424-3p/miR-450a-5p/miR-450a-5p/miR-450a-3p/miR-503/miR-542-5p/miR-542-3p/miR-450b-5p/miR-450b-3p	10
cl228_chr11	EPN_B	down	miR-34b-5p/miR-34b-3p/miR-34c-5p/miR-34c-3p	4
cl347_chr14	EPN_B	up	miR-127-5p/miR-127-3p/miR-136-5p/miR-136-3p/miR-337-5p/miR-337-3p/miR-431-5p/miR-431-3p/miR-433/miR-493-5p/miR-493-3p/miR-432-5p/miR-432-3p/miR-665	14
cl349_chr14	EPN_B	up	miR-134/miR-154-5p/miR-154-3p/miR-299-5p/miR-299-3p/miR-376c/miR-369-5p/miR-369-3p/miR-376a-5p/miR-376a-3p/miR-377-5p/miR-377-3p/miR-379-5p/miR-379-3p/miR-380-5p/miR-380-3p/miR-381/miR-382-5p/miR-382-3p/miR-323a-5p/miR-323a-3p/miR-329/miR-329/miR-409-5p/miR-409-3p/miR-412/miR-410/miR-376b/miR-485-5p/miR-485-3p/miR-487a/miR-494/	60

			miR-495/miR-496/miR-539-5p/miR-539-3p/miR-544a/miR-376a-3p/miR-487b/miR-411-5p/miR-411-3p/miR-654-5p/miR-654-3p/miR-655/miR-656/miR-758/miR-668/miR-1185-5p/miR-1185-2-3p/miR-1185-5p/miR-1185-1-3p/miR-300/miR-541-5p/miR-541-3p/miR-889/miR-543/miR-1197/miR-1193/miR-323b-5p/miR-323b-3p	
cl588_chr19	EPN_B	down	let-7e-5p/let-7e-3p/miR-125a-5p/miR-125a-3p/miR-99b-5p/miR-99b-3p	6
cl590_chr19	EPN_B	down	miR-512-5p/miR-512-3p/miR-512-5p/miR-512-3p/miR-498/miR-520e/miR-515-5p/miR-515-3p/miR-519e-5p/miR-519e-3p/miR-520f/miR-515-5p/miR-515-3p/miR-519c-5p/miR-519c-3p/miR-520a-5p/miR-520a-3p/miR-526b-5p/miR-526b-3p/miR-519b-5p/miR-519b-3p/miR-525-5p/miR-525-3p/miR-523-5p/miR-523-3p/miR-518f-5p/miR-518f-3p/miR-520b/miR-518b/miR-526a/miR-520c-5p/miR-520c-3p/miR-518c-5p/miR-518c-3p/miR-524-5p/miR-524-3p/miR-517-5p/miR-517a-3p/miR-519d/miR-521/miR-520d-5p/miR-520d-3p/miR-517b-3p/miR-517-5p/miR-520 g/miR-516b-5p/miR-516b-3p/miR-526a/miR-518e-5p/miR-518e-3p/miR-518a-5p/miR-518a-3p/miR-518d-5p/miR-518d-3p/miR-516b-5p/miR-516b-3p/miR-518a-5p/miR-518a-3p/miR-517-5p/miR-517c-3p/miR-520 h/miR-521/miR-522-5p/miR-522-3p/miR-519a-5p/miR-519a-3p/miR-527/miR-516a-5p/miR-516a-3p/miR-516a-5p/miR-516a-3p/miR-519a-3p/miR-1323/miR-1283/miR-1283	75
cl725_chr21	EPN_B	down	let-7c/miR-99a-5p/miR-99a-3p	3
cl81_chr1	EPN_B	up	miR-199a-5p/miR-199a-3p/miR-214-5p/miR-214-3p/miR-3120-3p/miR-3120-5p	6
cl1189_chrX	EPN_C	up	miR-188-5p/miR-188-3p/miR-362-5p/miR-362-3p/miR-500a-5p/miR-500a-3p/miR-501-5p/miR-501-3p/miR-502-5p/miR-502-3p/miR-532-5p/miR-532-3p/miR-660-5p/miR-660-3p/miR-500b	15
cl349_chr14	EPN_E	down	miR-134/miR-154-5p/miR-154-3p/miR-299-5p/miR-299-3p/miR-376c/miR-369-5p/miR-369-3p/miR-376a-5p/miR-376a-3p/miR-377-5p/miR-377-3p/miR-379-5p/miR-379-3p/miR-380-5p/miR-380-3p/miR-381/miR-382-5p/miR-382-3p/miR-323a-5p/miR-323a-3p/miR-329/miR-329/miR-409-5p/miR-409-3p/miR-412/miR-410/miR-376b/miR-485-5p/miR-485-3p/miR-487a/miR-494/miR-495/miR-496/miR-539-5p/miR-539-3p/miR-544a/miR-376a-3p/miR-487b/miR-411-5p/miR-411-3p/miR-654-5p/miR-654-3p/miR-655/miR-656/miR-758/miR-668/miR-1185-5p/miR-1185-2-3p/miR-1185-5p/miR-1185-1-3p/miR-300/miR-541-5p/miR-541-3p/miR-889/miR-543/miR-1197/miR-1193/miR-323b-5p/miR-323b-3p	60
cl590_chr19	EPN_F	up	miR-512-5p/miR-512-3p/miR-512-5p/miR-512-3p/miR-498/miR-520e/miR-515-5p/miR-515-3p/miR-519e-5p/miR-519e-3p/miR-520f/miR-515-5p/miR-515-3p/miR-519c-5p/miR-519c-3p/miR-520a-5p/miR-520a-3p/miR-526b-5p/miR-526b-3p/miR-519b-5p/miR-519b-3p/miR-525-5p/miR-	75

(continued)

Table 10.1 (continued)

Cluster ID	Expr. class	Cluster sign	Cluster members	Size
			525-3p/miR-523-5p/miR-523-3p/miR-518f-5p/miR-518f-3p/miR-520b/miR-518b/miR-526a/miR-520c-5p/miR-520c-3p/miR-518c-5p/miR-518c-3p/miR-524-5p/miR-524-3p/miR-517-5p/miR-517a-3p/miR-519d/miR-521/miR-520d-5p/miR-520d-3p/miR-517b-3p/miR-517-5p/miR-520 g/miR-516b-5p/miR-516b-3p/miR-526a/miR-518e-5p/miR-518e-3p/miR-518a-5p/miR-518a-3p/miR-518d-5p/miR-518d-3p/miR-516b-5p/miR-516b-3p/miR-518a-5p/miR-518a-3p/miR-517-5p/miR-517c-3p/miR-520 h/miR-521/miR-522-5p/miR-522-3p/miR-519a-5p/miR-519a-3p/miR-527/miR-516a-5p/miR-516a-3p/miR-516a-5p/miR-516a-3p/miR-519a-3p/miR-1323/miR-1283/miR-1283	
cl349_chr14	EPN_G	down	miR-134/miR-154-5p/miR-154-3p/miR-299-5p/miR-299-3p/miR-376c/miR-369-5p/miR-369-3p/miR-376a-5p/miR-376a-3p/miR-377-5p/miR-377-3p/miR-379-5p/miR-379-3p/miR-380-5p/miR-380-3p/miR-381/miR-382-5p/miR-382-3p/miR-323a-5p/miR-323a-3p/miR-329/miR-329/miR-409-5p/miR-409-3p/miR-412/miR-410/miR-376b/miR-485-5p/miR-485-3p/miR-487a/miR-494/miR-495/miR-496/miR-539-5p/miR-539-3p/miR-544a/miR-376a-3p/miR-487b/miR-411-5p/miR-411-3p/miR-654-5p/miR-654-3p/miR-655/miR-656/miR-758/miR-668/miR-1185-5p/miR-1185-2-3p/miR-1185-5p/miR-1185-1-3p/miR-300/miR-541-5p/miR-541-3p/miR-889/miR-543/miR-1197/miR-1193/miR-323b-5p/miR-323b-3p	60
cl1189_chrX	EPN_H	down	miR-188-5p/miR-188-3p/miR-362-5p/miR-362-3p/miR-500a-5p/miR-500a-3p/miR-501-5p/miR-501-3p/miR-502-5p/miR-502-3p/miR-532-5p/miR-532-3p/miR-660-5p/miR-660-3p/miR-500b	15
cl1189_chrX	EPN_I	down	miR-188-5p/miR-188-3p/miR-362-5p/miR-362-3p/miR-500a-5p/miR-500a-3p/miR-501-5p/miR-501-3p/miR-502-5p/miR-502-3p/miR-532-5p/miR-532-3p/miR-660-5p/miR-660-3p/miR-500b	15
cl1195_chrX	EPN_I	up	miR-421/miR-374b-5p/miR-374b-3p/miR-374c-5p/miR-374c-3p	5
cl228_chr11	EPN_I	down	miR-34b-5p/miR-34b-3p/miR-34c-5p/miR-34c-3p	4
cl349_chr14	EPN_I	down	miR-134/miR-154-5p/miR-154-3p/miR-299-5p/miR-299-3p/miR-376c/miR-369-5p/miR-369-3p/miR-376a-5p/miR-376a-3p/miR-377-5p/miR-377-3p/miR-379-5p/miR-379-3p/miR-380-5p/miR-380-3p/miR-381/miR-382-5p/miR-382-3p/miR-323a-5p/miR-323a-3p/miR-329/miR-329/miR-409-5p/miR-409-3p/miR-412/miR-410/miR-376b/miR-485-5p/miR-485-3p/miR-487a/miR-494/miR-495/miR-496/miR-539-5p/miR-539-3p/miR-544a/miR-376a-3p/miR-487b/miR-411-5p/miR-411-3p/miR-654-5p/miR-654-3p/miR-655/miR-656/miR-758/miR-668/miR-1185-5p/miR-1185-2-3p/miR-1185-5p/miR-1185-1-3p/miR-300/miR-541-5p/miR-541-3p/miR-889/miR-543/miR-1197/miR-1193/miR-323b-5p/miR-323b-3p	60

cl349_chr14	EPN_I	up	miR-134/miR-154-5p/miR-154-3p/miR-299-5p/miR-299-3p/miR-376c/miR-369-5p/miR-369-3p/miR-376a-5p/miR-376a-3p/miR-377-5p/miR-377-3p/miR-379-5p/miR-379-3p/miR-380-5p/miR-380-3p/miR-381/miR-382-5p/miR-382-3p/miR-323a-5p/miR-323a-3p/miR-329/miR-329/miR-409-5p/miR-409-3p/miR-412/miR-410/miR-376b/miR-485-5p/miR-485-3p/miR-487a/miR-494/miR-495/miR-496/miR-539-5p/miR-539-3p/miR-544a/miR-376a-3p/miR-487b/miR-411-5p/miR-411-3p/miR-654-5p/miR-654-3p/miR-655/miR-656/miR-758/miR-668/miR-1185-5p/miR-1185-2-3p/miR-1185-5p/miR-1185-1-3p/miR-300/miR-541-5p/miR-541-3p/miR-889/miR-543/miR-1197/miR-1193/miR-323b-5p/miR-323b-3p	60
cl590_chr19	EPN_I	down	miR-512-5p/miR-512-3p/miR-512-5p/miR-512-3p/miR-498/miR-520e/miR-515-5p/miR-515-3p/miR-519e-5p/miR-519e-3p/miR-520f/miR-515-5p/miR-515-3p/miR-519c-5p/miR-519c-3p/miR-520a-5p/miR-520a-3p/miR-526b-5p/miR-526b-3p/miR-519b-5p/miR-519b-3p/miR-525-5p/miR-525-3p/miR-523-5p/miR-523-3p/miR-518f-5p/miR-518f-3p/miR-520b/miR-518b/miR-526a/miR-520c-5p/miR-520c-3p/miR-518c-5p/miR-518c-3p/miR-524-5p/miR-524-3p/miR-517-5p/miR-517a-3p/miR-519d/miR-521/miR-520d-5p/miR-520d-3p/miR-517b-3p/miR-517-5p/miR-520 g/miR-516b-5p/miR-516b-3p/miR-526a/miR-518e-5p/miR-518e-3p/miR-518a-5p/miR-518a-3p/miR-518d-5p/miR-518d-3p/miR-516b-5p/miR-516b-3p/miR-518a-5p/miR-518a-3p/miR-517-5p/miR-517c-3p/miR-520 h/miR-521/miR-522-5p/miR-522-3p/miR-519a-5p/miR-519a-3p/miR-527/miR-516a-5p/miR-516a-3p/miR-516a-5p/miR-516a-3p/miR-519a-3p/miR-1323/miR-1283/miR-1283	75
cl1190_chrX	MB_WNT	up	let-7f-2-3p/let-7f-5p/miR-98	3
cl1226_chrX	MB_WNT	down	miR-513a-5p/miR-513a-3p/miR-506-3p/miR-506-5p/miR-507/miR-508-5p/miR-508-3p	7
cl1227_chrX	MB_WNT	down	miR-509-5p/miR-509-3p/miR-509-5p/miR-509-3p/miR-509-3-5p/miR-509-3p/miR-514b-5p/miR-514b-3p	8
cl2_chr1	MB_WNT	up	miR-200b-5p/miR-200b-3p/miR-200a-5p/miR-200a-3p/miR-429	5
cl904_chr5	MB_WNT	up	miR-449a/miR-449b-5p/miR-449b-3p/miR-449c-5p/miR-449c-3p	5
cl96_chr1	MB_WNT	up	miR-215/miR-194-5p	2
cl1161_chr9	MB_SHH	up	miR-199b-5p/miR-199b-3p/miR-3154	3
cl267_chr12	MB_SHH	down	miR-331-5p/miR-331-3p/miR-3685	3
cl512_chr17	MB_SHH	down	miR-338-5p/miR-338-3p/miR-657/miR-1250/miR-3065-5p/miR-3065-3p	6
cl564_chr19	MB_SHH	down	miR-181c-5p/miR-181c-3p/miR-181d	3

(continued)

Table 10.1 (continued)

Cluster ID	Expr. class	Cluster sign	Cluster members	Size
cl742_chr22	MB_SHH	up	miR-130b-5p/miR-130b-3p/miR-301b	3
cl81_chr1	MB_SHH	up	miR-199a-5p/miR-199a-3p/miR-214-5p/miR-214-3p/miR-3120-3p/miR-3120-5p	6
cl904_chr5	MB_SHH	up	miR-449a/miR-449b-5p/miR-449b-3p/miR-449c-5p/miR-449c-3p	5
cl1195_chrX	MB_G3	down	miR-421/miR-374b-5p/miR-374b-3p/miR-374c-5p/miR-374c-3p	5
cl1222_chrX	MB_G3	down	miR-892a/miR-890/miR-891b/miR-888-5p/miR-888-3p/miR-892b	6
cl347_chr14	MB_G3	down	miR-127-5p/miR-127-3p/miR-136-5p/miR-136-3p/miR-337-5p/miR-337-3p/miR-431-5p/miR-431-3p/miR-433/miR-493-5p/miR-493-3p/miR-432-5p/miR-432-3p/miR-665	14
cl349_chr14	MB_G3	down	miR-134/miR-154-5p/miR-154-3p/miR-299-5p/miR-299-3p/miR-376c/miR-369-5p/miR-369-3p/miR-376a-5p/miR-376a-3p/miR-377-5p/miR-377-3p/miR-379-5p/miR-379-3p/miR-380-5p/miR-380-3p/miR-381/miR-382-5p/miR-382-3p/miR-323a-5p/miR-323a-3p/miR-329/miR-329/miR-409-5p/miR-409-3p/miR-412/miR-410/miR-376b/miR-485-5p/miR-485-3p/miR-487a/miR-494/miR-495/miR-496/miR-539-5p/miR-539-3p/miR-544a/miR-376a-3p/miR-487b/miR-411-5p/miR-411-3p/miR-654-5p/miR-654-3p/miR-655/miR-656/miR-758/miR-668/miR-1185-5p/miR-1185-2-3p/miR-1185-5p/miR-1185-1-3p/miR-300/miR-541-5p/miR-541-3p/miR-889/miR-543/miR-1197/miR-1193/miR-323b-5p/miR-323b-3p	60
cl51_chr1	MB_G3	up	miR-137/miR-2682-5p/miR-2682-3p	3
cl522_chr18	MB_G3	up	miR-1/miR-133a	2
cl588_chr19	MB_G3	up	let-7e-5p/let-7e-3p/miR-125a-5p/miR-125a-3p/miR-99b-5p/miR-99b-3p	6
cl591_chr19	MB_G3	up	miR-371a-5p/miR-371a-3p/miR-372/miR-373-5p/miR-373-3p/miR-371b-5p/miR-371b-3p	7
cl1118_chr8	MB_G4	down	miR-939/miR-1234	2
cl1161_chr9	MB_G4	down	miR-199b-5p/miR-199b-3p/miR-3154	3
cl1217_chrX	MB_G4	down	miR-424-5p/miR-424-3p/miR-450a-5p/miR-450a-5p/miR-450a-3p/miR-503/miR-542-5p/miR-542-3p/miR-450b-5p/miR-450b-3p	10
cl1227_chrX	MB_G4	up	miR-509-5p/miR-509-3p/miR-509-5p/miR-509-3p/miR-509-3-5p/miR-509-3p/miR-514b-5p/miR-514b-3p	8

cl349_chr14	MB_G4	up	miR-134/miR-154-5p/miR-154-3p/miR-299-5p/miR-299-3p/miR-376c/miR-369-5p/miR-369-3p/miR-376a-5p/miR-376a-3p/miR-377-5p/miR-377-3p/miR-379-5p/miR-379-3p/miR-380-5p/miR-380-3p/miR-381/miR-382-5p/miR-382-3p/miR-323a-5p/miR-323a-3p/miR-329/miR-329/miR-409-5p/miR-409-3p/miR-412/miR-410/miR-376b/miR-485-5p/miR-485-3p/miR-487a/miR-494/miR-495/miR-496/miR-539-5p/miR-539-3p/miR-544a/miR-376a-3p/miR-487b/miR-411-5p/miR-411-3p/miR-654-5p/miR-654-3p/miR-655/miR-656/miR-758/miR-668/miR-1185-5p/miR-1185-2-3p/miR-1185-5p/miR-1185-1-3p/miR-300/miR-541-5p/miR-541-3p/miR-889/miR-543/miR-1197/miR-1193/miR-323b-5p/miR-323b-3p	60
cl512_chr17	MB_G4	up	miR-338-5p/miR-338-3p/miR-657/miR-1250/miR-3065-5p/miR-3065-3p	6
cl590_chr19	MB_G4	up	miR-512-5p/miR-512-3p/miR-512-5p/miR-512-3p/miR-498/miR-520e/miR-515-5p/miR-515-3p/miR-519e-5p/miR-519e-3p/miR-520f/miR-515-5p/miR-515-3p/miR-519c-5p/miR-519c-3p/miR-520a-5p/miR-520a-3p/miR-526b-5p/miR-526b-3p/miR-519b-5p/miR-519b-3p/miR-525-5p/miR-525-3p/miR-523-5p/miR-523-3p/miR-518f-5p/miR-518f-3p/miR-520b/miR-518b/miR-526a/miR-520c-5p/miR-520c-3p/miR-518c-5p/miR-518c-3p/miR-524-5p/miR-524-3p/miR-517-5p/miR-517a-3p/miR-519d/miR-521/miR-520d-5p/miR-520d-3p/miR-517b-3p/miR-517-5p/miR-520 g/miR-516b-5p/miR-516b-3p/miR-526a/miR-518e-5p/miR-518e-3p/miR-518a-5p/miR-518a-3p/miR-518d-5p/miR-518d-3p/miR-516b-5p/miR-516b-3p/miR-518a-5p/miR-518a-3p/miR-517-5p/miR-517c-3p/miR-520 h/miR-521/miR-522-5p/miR-522-3p/miR-519a-5p/miR-519a-3p/miR-527/miR-516a-5p/miR-516a-3p/miR-516a-5p/miR-516a-3p/miR-519a-3p/miR-1323/miR-1283/miR-1283	75
cl742_chr22	MB_G4	down	miR-130b-5p/miR-130b-3p/miR-301b	3
cl81_chr1	MB_G4	down	miR-199a-5p/miR-199a-3p/miR-214-5p/miR-214-3p/miR-3120-3p/miR-3120-5p	6
cl1217_chrX	NB_noMYCN	up	miR-424-5p/miR-424-3p/miR-450a-5p/miR-450a-5p/miR-450a-3p/miR-503/miR-542-5p/miR-542-3p/miR-450b-5p/miR-450b-3p	10
cl347_chr14	NB_noMYCN	up	miR-127-5p/miR-127-3p/miR-136-5p/miR-136-3p/miR-337-5p/miR-337-3p/miR-431-5p/miR-431-3p/miR-433/miR-493-5p/miR-493-3p/miR-432-5p/miR-432-3p/miR-665	14
cl349_chr14	NB_noMYCN	up	miR-134/miR-154-5p/miR-154-3p/miR-299-5p/miR-299-3p/miR-376c/miR-369-5p/miR-369-3p/miR-376a-5p/miR-376a-3p/miR-377-5p/miR-377-3p/miR-379-5p/miR-379-3p/miR-380-5p/miR-380-3p/miR-381/miR-382-5p/miR-382-3p/miR-323a-5p/miR-323a-3p/miR-329/miR-329/miR-409-5p/miR-409-3p/miR-412/miR-410/miR-376b/miR-485-5p/miR-485-3p/miR-487a/miR-494/miR-495/miR-496/miR-539-5p/miR-539-3p/miR-544a/miR-376a-3p/miR-487b/miR-411-5p/	60

(continued)

Table 10.1 (continued)

Cluster ID	Expr. class	Cluster sign	Cluster members	Size
			miR-411-3p/miR-654-5p/miR-654-3p/miR-655/miR-656/miR-758/miR-668/miR-1185-5p/miR-1185-2-3p/miR-1185-5p/miR-1185-1-3p/miR-300/miR-541-5p/miR-541-3p/miR-889/miR-543/miR-1197/miR-1193/miR-323b-5p/miR-323b-3p	
cl590_chr19	NB_noMYCN	down	miR-512-5p/miR-512-3p/miR-512-5p/miR-512-3p/miR-498/miR-520e/miR-515-5p/miR-515-3p/miR-519e-5p/miR-519e-3p/miR-520f/miR-515-5p/miR-515-3p/miR-519c-5p/miR-519c-3p/miR-520a-5p/miR-520a-3p/miR-526b-5p/miR-526b-3p/miR-519b-5p/miR-519b-3p/miR-525-5p/miR-525-3p/miR-523-5p/miR-523-3p/miR-518f-5p/miR-518f-3p/miR-520b/miR-518b/miR-526a/miR-520c-5p/miR-520c-3p/miR-518c-5p/miR-518c-3p/miR-524-5p/miR-524-3p/miR-517-5p/miR-517a-3p/miR-519d/miR-521/miR-520d-5p/miR-520d-3p/miR-517b-3p/miR-517-5p/miR-520 g/miR-516b-5p/miR-516b-3p/miR-526a/miR-518e-5p/miR-518e-3p/miR-518a-5p/miR-518a-3p/miR-518d-5p/miR-518d-3p/miR-516b-5p/miR-516b-3p/miR-518a-5p/miR-518a-3p/miR-517-5p/miR-517c-3p/miR-520 h/miR-521/miR-522-5p/miR-522-3p/miR-519a-5p/miR-519a-3p/miR-527/miR-516a-5p/miR-516a-3p/miR-516a-5p/miR-516a-3p/miR-519a-3p/miR-1323/miR-1283/miR-1283	75
cl590_chr19	NB_noMYCN	up	miR-512-5p/miR-512-3p/miR-512-5p/miR-512-3p/miR-498/miR-520e/miR-515-5p/miR-515-3p/miR-519e-5p/miR-519e-3p/miR-520f/miR-515-5p/miR-515-3p/miR-519c-5p/miR-519c-3p/miR-520a-5p/miR-520a-3p/miR-526b-5p/miR-526b-3p/miR-519b-5p/miR-519b-3p/miR-525-5p/miR-525-3p/miR-523-5p/miR-523-3p/miR-518f-5p/miR-518f-3p/miR-520b/miR-518b/miR-526a/miR-520c-5p/miR-520c-3p/miR-518c-5p/miR-518c-3p/miR-524-5p/miR-524-3p/miR-517-5p/miR-517a-3p/miR-519d/miR-521/miR-520d-5p/miR-520d-3p/miR-517b-3p/miR-517-5p/miR-520 g/miR-516b-5p/miR-516b-3p/miR-526a/miR-518e-5p/miR-518e-3p/miR-518a-5p/miR-518a-3p/miR-518d-5p/miR-518d-3p/miR-516b-5p/miR-516b-3p/miR-518a-5p/miR-518a-3p/miR-517-5p/miR-517c-3p/miR-520 h/miR-521/miR-522-5p/miR-522-3p/miR-519a-5p/miR-519a-3p/miR-527/miR-516a-5p/miR-516a-3p/miR-516a-5p/miR-516a-3p/miR-519a-3p/miR-1323/miR-1283/miR-1283	75
cl81_chr1	NB_noMYCN	down	miR-199a-5p/miR-199a-3p/miR-214-5p/miR-214-3p/miR-3120-3p/miR-3120-5p	6
cl1217_chrX	NB_MYCN	down	miR-424-5p/miR-424-3p/miR-450a-5p/miR-450a-5p/miR-450a-3p/miR-503/miR-542-5p/miR-542-3p/miR-450b-5p/miR-450b-3p	10
cl1226_chrX	NB_MYCN	down	miR-513a-5p/miR-513a-3p/miR-506-3p/miR-506-5p/miR-507/miR-508-5p/miR-508-3p	7

cl349_chr14	NB_MYCN	down	miR-134/miR-154-5p/miR-154-3p/miR-299-5p/miR-299-3p/miR-376c/miR-369-5p/miR-369-3p/miR-376a-5p/miR-376a-3p/miR-377-5p/miR-377-3p/miR-379-5p/miR-379-3p/miR-380-5p/miR-380-3p/miR-381/miR-382-5p/miR-382-3p/miR-323a-5p/miR-323a-3p/miR-329/miR-329/miR-409-5p/miR-409-3p/miR-412/miR-410/miR-376b/miR-485-5p/miR-485-3p/miR-487a/miR-494/miR-495/miR-496/miR-539-5p/miR-539-3p/miR-544a/miR-376a-3p/miR-487b/miR-411-5p/miR-411-3p/miR-654-5p/miR-654-3p/miR-655/miR-656/miR-758/miR-668/miR-1185-5p/miR-1185-2-3p/miR-1185-5p/miR-1185-1-3p/miR-300/miR-541-5p/miR-541-3p/miR-889/miR-543/miR-1197/miR-1193/miR-323b-5p/miR-323b-3p	60
cl590_chr19	NB_MYCN	down	miR-512-5p/miR-512-3p/miR-512-5p/miR-512-3p/miR-498/miR-520e/miR-515-5p/miR-515-3p/miR-519e-5p/miR-519e-3p/miR-520f/miR-515-5p/miR-515-3p/miR-519c-5p/miR-519c-3p/miR-520a-5p/miR-520a-3p/miR-526b-5p/miR-526b-3p/miR-519b-5p/miR-519b-3p/miR-525-5p/miR-525-3p/miR-523-5p/miR-523-3p/miR-518f-5p/miR-518f-3p/miR-520b/miR-518b/miR-526a/miR-520c-5p/miR-520c-3p/miR-518c-5p/miR-518c-3p/miR-524-5p/miR-524-3p/miR-517-5p/miR-517a-3p/miR-519d/miR-521/miR-520d-5p/miR-520d-3p/miR-517b-3p/miR-517-5p/miR-520 g/miR-516b-5p/miR-516b-3p/miR-526a/miR-518e-5p/miR-518e-3p/miR-518a-5p/miR-518a-3p/miR-518d-5p/miR-518d-3p/miR-516b-5p/miR-516b-3p/miR-518a-5p/miR-518a-3p/miR-517-5p/miR-517c-3p/miR-520 h/miR-521/miR-522-5p/miR-522-3p/miR-519a-5p/miR-519a-3p/miR-527/miR-516a-5p/miR-516a-3p/miR-516a-5p/miR-516a-3p/miR-519a-3p/miR-1323/miR-1283/miR-1283	75
cl590_chr19	NB_MYCN	up	miR-512-5p/miR-512-3p/miR-512-5p/miR-512-3p/miR-498/miR-520e/miR-515-5p/miR-515-3p/miR-519e-5p/miR-519e-3p/miR-520f/miR-515-5p/miR-515-3p/miR-519c-5p/miR-519c-3p/miR-520a-5p/miR-520a-3p/miR-526b-5p/miR-526b-3p/miR-519b-5p/miR-519b-3p/miR-525-5p/miR-525-3p/miR-523-5p/miR-523-3p/miR-518f-5p/miR-518f-3p/miR-520b/miR-518b/miR-526a/miR-520c-5p/miR-520c-3p/miR-518c-5p/miR-518c-3p/miR-524-5p/miR-524-3p/miR-517-5p/miR-517a-3p/miR-519d/miR-521/miR-520d-5p/miR-520d-3p/miR-517b-3p/miR-517-5p/miR-520 g/miR-516b-5p/miR-516b-3p/miR-526a/miR-518e-5p/miR-518e-3p/miR-518a-5p/miR-518a-3p/miR-518d-5p/miR-518d-3p/miR-516b-5p/miR-516b-3p/miR-518a-5p/miR-518a-3p/miR-517-5p/miR-517c-3p/miR-520 h/miR-521/miR-522-5p/miR-522-3p/miR-519a-5p/miR-519a-3p/miR-527/miR-516a-5p/miR-516a-3p/miR-516a-5p/miR-516a-3p/miR-519a-3p/miR-1323/miR-1283/miR-1283	75

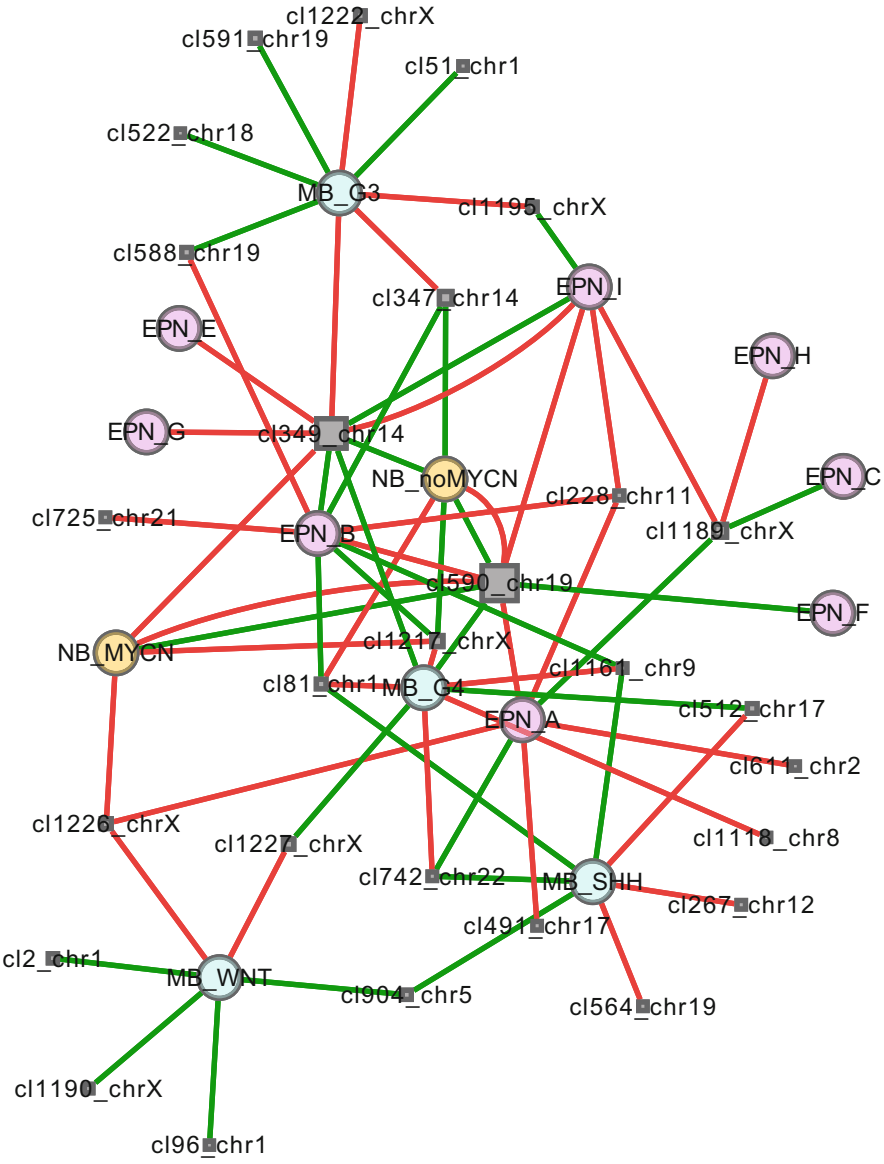


Fig. 10.2 Results of the clustMMRA pipeline applied to three pediatric cancers, namely NB, MB, and EPN. The bipartite graph represents two sets of nodes (miRNA clusters and cancer subgroups) connected by links corresponding to the potential regulation predicted by clustMMRA. Yellow, blue, and pink nodes correspond, respectively, to NB, MB, and EPN subgroups, gray nodes correspond to miRNA clusters. The gray nodes correspond to the miRNA clusters, with the size of the node proportional to the number of miRNAs belonging to the cluster. Edges are green when the miRNA cluster is upregulated in a given subgroup, red when it is downregulated

in EPN and the MYCN amplified and not amplified subgroups in NB. Interestingly, the cluster of miRNAs at chromosome 14q32 was already found to be associated to relapse in ependymoma (Costa et al. 2011), downregulated in desmoplastic MB (Lucon et al. 2013a) and associated with poor prognosis in NB (Roth et al. 2016).

Some disease subgroups show connections with a high number of clusters, in particular, the G3 and G4 subgroups in MB are linked to eight and nine clusters, respectively, and the EPN subgroup B which is linked to nine clusters. Specifically in G3 and G4 of MB, mRNA expression has been shown to correlate poorly with protein expression (Forget et al. 2018), suggesting a hallmark role of posttranscriptional regulation in driving these cancer subgroups, possibly involving miRNA regulation. The WNT and G3 subgroups in MB show more specific miRNA regulation than the other subgroups, as they have specific links to specific clusters (3 and 4, respectively) that are not linked to any other subgroup or other cancers.

10.5 Discussion

In the past decades, we have observed a tremendous increase in the availability of omics profiles from normal to cancer samples. For example, The Cancer Genome Atlas (TCGA), a landmark cancer genomics program, molecularly characterized over 20,000 primary cancer and matched normal samples spanning 33 cancer types. The comprehensive genomics data generated by TCGA includes the unprecedented amount of ~11.000 libraries of miRNA sequences (Chu et al. 2016). To support the analysis of such an amount of data, scalability of bioinformatics pipelines in terms of CPU usage, data storage, and memory requirements is increasingly important to handle larger datasets.

Here, we present a scalable implementation of the clustMMRA pipeline for the unbiased identification of clusters of miRNAs that potentially regulate cancer subtyping. This software allows to highlight the regulation by multiple miRNA species that act cooperatively on the phenotypes of cancer subgroups. This is quite original, since the majority of computational and experimental approaches for the identification of master miRNA regulators involved in cancer onset and subtyping are typically designed to detect the regulatory effect of a single miRNA.

The new version of the clustMMRA pipeline shows some advantages over the previous one. Specifically, the first version of clustMMRA had to run on a cluster, due to its computing needs in terms of both CPU power and memory usage. This update removed the reliance on a supporting computing cluster, as ClustMMRA can now run independently on a laptop with a recent processor and 16 Gb RAM. In this new version of clustMMRA, the third-party ARACNE software has been updated to the latest version, achieving a dramatic improvement in computational performance in both time and memory usage while preserving the network inference accuracy of the original algorithm. Finally, the readability of the code and its documentation have been significantly improved.

As a case study, we have applied the clustMMRA pipeline to multiple pediatric cancers to highlight genomic clusters of miRNAs potentially involved in several subgroups of the different cancers or specifically involved in the phenotype of a subgroup. Three datasets ranging in size up to a hundred samples were analyzed, and the entire analysis process was done in a few minutes. The scalability of the software to large datasets is therefore very satisfactory.

The results of our analysis confirm a possible regulatory role of the C14MC cluster at the 14q32 locus in several pediatric cancers. This cluster is downregulated in subgroups with aggressive phenotype, in particular, in the MB G3 subgroup and the NB MYCN amplified subgroup, while it is upregulated in the NB non-amplified MYCN subgroup. These observations are consistent with previous studies showing a tumor suppressor role in the C14MC cluster.

In our results, we did not observe cluster miR-17/92 as a regulator of any subgroup, as we would have expected from the wide literature showing its overexpression in a variety of human cancers, including pediatric cancers. This may be due to the fact that the miR-17/92 cluster is very often dysregulated in human cancers, but its dysregulation could occur in all cancers and not be specific to a subgroup of the diseases we have analyzed.

Looking in more detail at the results for this cluster, we observe that miR-17/92 is selected in step (i) of the pipeline, as it is differentially expressed in some cancer subgroups, but is discarded at step (ii) because the signatures of these subgroups are not enriched by the targets of this cluster. Indeed, results of clustMMRA rely on well-defined transcriptomic signatures specific to subgroups. If these signatures are not accurate enough or are predominantly associated with regulatory factors other than miRNAs, some potentially relevant clusters are not retained in the clustMMRA output. With the availability of large-scale quantitative proteomics data, the same analysis can be performed using signatures specific to subgroups defined based on both transcriptomics and proteomics profiles.

Finally, in our case study, clustMMRA was applied to systematically investigate the cooperative effect of miRNAs belonging to genomic clusters. In practice, the use of clustMMRA can be generalized in order to study other sets of cooperatively acting miRNAs than the case of genomic clusters, such as co-expressed miRNAs from different genomic locations.

10.6 Code Availability

The clustMMRA pipeline is available at https://github.com/sysbio-curie/clustMMRA_v2.

Acknowledgments This work was supported by grants from ITMO Cancer AVIESAN (National Alliance for Life Sciences and Health), within the framework of the Plan Cancer 2014–2019 and convention "2018, Non-coding RNA in cancerology: fundamental to translational (18CN039-00)" and the European Commission's Horizon 2020 Program, H2020-SC1-DTH-2018-1, "iPC—individualizedPaediatricCure" (ref. 826121).

References

Agarwal V, Bell GW, Nam JW, Bartel DP (2015) Predicting effective microRNA target sites in mammalian mRNAs. elife 4:e05005

Altuvia Y, Landgraf P, Lithwick G, Elefant N, Pfeffer S, Aravin A, Brownstein MJ et al (2005) Clustering and conservation patterns of human microRNAs. Nucleic Acids Res 33(8): 2697–2706

Anders G, Mackowiak SD, Jens M, Maaskola J, Kuntzagk A, Rajewsky N,Landthaler M, et al (2012) doRiNA: a database of RNA interactions in post-transcriptional regulation. Nucleic Acids Res (Database issue)

Baek D, Villén J, Shin C, Camargo FD, Gygi SP, Bartel DP (2008) The impact of microRNAs on protein output. Nature 455(7209):64–71

Bartel DP (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 116(2): 281–297

Bartel DP (2018) Metazoan MicroRNAs. Cell 173(1):20–51

Bhavsar SP, Løkke C, Flægstad T, Einvik C (2018) Hsa-miR-376c-3p targets Cyclin D1 and induces G1-cell cycle arrest in neuroblastoma cells. Oncol Lett 16(5):6786–6794

Brabletz S, Brabletz T (2010) The ZEB/miR-200 feedback loop–a motor of cellular plasticity in development and cancer? EMBO Rep 11:670–677

Calin GA, Croce CM (2006) MicroRNA signatures in human cancers. Nat Rev Cancer 6(11): 857–866

Cantini L, Bertoli G, Cava C, Dubois T, Zinovyev A, Caselle M, Castiglioni I et al (2019) Identification of microRNA clusters cooperatively acting on epithelial to mesenchymal transition in triple negative breast cancer. Nucleic Acids Res 47(5):2205–2215

Chu A, Robertson G, Brooks D, Mungall AJ, Birol I, Coope R, Ma Y et al (2016) Large-scale profiling of microRNAs for The Cancer Genome Atlas. Nucleic Acids Res 44(1):e3

Cimmino A, Calin GA, Fabbri M, Iorio MV, Ferracin M, Shimizu M, Wojcik SE et al (2005) miR-15 and miR-16 induce apoptosis by targeting BCL2. Proc Natl Acad Sci U S A 102(39): 13944–13949

Costa FF, Bischof JM, Vanin EF, Lulla RR, Wang M, Sredni ST, Rajaram V et al (2011) Identification of microRNAs as potential prognostic markers in ependymoma. PLoS One 6(10):e25114

Cursons J, Pillman KA, Scheer KG, Gregory PA, Foroutan M, Hediyeh-Zadeh S, Toubia J et al (2018) Combinatorial Targeting by MicroRNAs Co-ordinates Post-transcriptional Control of EMT. Cell Syst 7(1):77–91.e7

Dexheimer PJ, Cochella L (2020) MicroRNAs: From Mechanism to Organism. Front Cell Dev Biol 8:409

Forget A, Martignetti L, Puget S, Calzone L, Brabetz S, Picard D, Montagud A et al (2018) Aberrant ERBB4-SRC Signaling as a Hallmark of Group 4 Medulloblastoma Revealed by Integrative Phosphoproteomic Profiling. Cancer Cell 34(3):379–395

Friedman RC, Farh KK, Burge CB, Bartel DP (2009) Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 19(1):92–105

Hausser J, Zavolan M (2014) Identification and consequences of miRNA–target interactions–beyond repression of gene expression. Nat Rev Genet 15:599–612

Hayashita Y, Osada H, Tatematsu Y, Yamada H, Yanagisawa K, Tomida S, Yatabe Y et al (2005) A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. Cancer Res 65(21):9628–9632

He L, Thomson JM, Hemann MT, Hernando-Monge E, Mu D, Goodson S, Powers S et al (2005) A microRNA polycistron as a potential human oncogene. Nature 435(7043):828–833

He J, Zhou Z, Reed M, Califano A (2017) Accelerated parallel algorithm for gene network reverse engineering. BMC Syst Biol 11:83

Hendrickson DG, Hogan DJ, McCullough HL, Myers JW, Herschlag D, Ferrell JE, Brown PO (2009) Concordant regulation of translation and mRNA abundance for hundreds of targets of a human microRNA. PLoS Biol 7(11):e1000238

Hill KE, Kelly AD, Kuijjer ML, Barry W, Rattani A, Garbutt CC, Kissick H et al (2017) An imprinted non-coding genomic cluster at 14q32 defines clinically relevant molecular subtypes in osteosarcoma across multiple independent datasets. J Hematol Oncol 10(1):107

Huang HY, Lin YC, Li J, Huang KY, Shrestha S, Hong HC, Tang Y et al (2020) miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. Nucleic Acids Res 48(D1):D148–D154

John CR, Watson D, Russ D, Goldmann K, Ehrenstein M, Pitzalis C, Barnes M (2020) M3C: Monte Carlo reference-based consensus clustering. Sci Rep 10(1):1–14

Johnson RA, Wright KD, Poppleton H, Mohankumar KM, Finkelstein D, Pounds SB et al (2010) Cross-species genomics matches driver mutations and cell compartments to model ependymoma. Nature 466(7306):632–636

Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E (2007) The role of site accessibility in microRNA target recognition. Nat Genet 39(10):1278–1284. https://doi.org/10.1038/ng2135

Kozomara A, Birgaoanu M, Griffiths-Jones S (2019) miRBase: from microRNA sequences to function. Nucleic Acids Res 47:D155–D162

Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P et al (2005) Combinatorial microRNA target predictions. Nat Genet 37(5):495–500

Lachmann A, Giorgi FM, Lopez G, Califano A (2016) ARACNe-AP: gene network reverse engineering through adaptive partitioning inference of mutual information. Bioinformatics 32(14):2233–2235

Liang H, Li WH (2007) MicroRNA regulation of human protein protein interaction network. RNA 13(9):1402–1408

Lim LP, Lau NC, Garrett-Engele P, Grimson A, Schelter JM, Castle J, Bartel DP et al (2005) Microarray analysis shows that some microRNAs downregulate large numbers of target mRNAs. Nature 433(7027):769–773

Liu W, Wang X (2019) Prediction of functional microRNA targets by integrative modeling of microRNA binding and target expression data. Genome Biol 20(1):18

Lucon DR, Rocha Cde S, Craveiro RB, Dilloo D, Cardinalli IA, Cavalcanti DP, Aguiar Sdos S et al (2013a) Downregulation of 14q32 microRNAs in Primary Human Desmoplastic Medulloblastoma. Front Oncol 25(3):254

Lucon DR, Rocha Cde S, Craveiro RB, Dilloo D, Cardinalli IA, Cavalcanti DP et al (2013bSep 25) Downregulation of 14q32 microRNAs in Primary Human Desmoplastic Medulloblastoma. Front Oncol 3:254

Ma L, Teruya-Feldstein J, Weinberg RA (2007) Tumour invasion and metastasis initiated by microRNA-10b in breast cancer. Nature 449:682–688

Marco A, Ninova M, Ronshaugen M, Griffiths-Jones S (2013) Clusters of microRNAs emerge by new hairpins in existing transcripts. Nucleic Acids Res 41(16):7745–7752

Margolin AA, Wang K, Lim WK, Kustagi M, Nemenman I, Califano A (2006) Reverse engineering cellular networks. Nat Protoc 1(2):662–671

Mestdagh P, Boström AK, Impens F, Fredlund E, Van Peer G, De Antonellis P, von Stedingk K et al (2010) The miR-17-92 microRNA cluster regulates multiple components of the TGF-β pathway in neuroblastoma. Mol Cell 40(5):762–773

Monti S, Tamayo P, Mesirov J, Golub T (2003) Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data. Mach Learn 52(1): 91–118

Mu P, Han YC, Betel D, Yao E, Squatrito M, Ogrodowski P, de Stanchina E et al (2009) Genetic dissection of the miR-17~92 cluster of microRNAs in Myc-induced B-cell lymphomas. Genes Dev 23(24):2806–2811

Murphy BL, Obad S, Bihannic L, Ayrault O, Zindy F, Kauppinen S, Roussel MF (2013) Silencing of the miR-17~92 cluster family inhibits medulloblastoma progression. CancerRes 73(23): 7068–7078

Olive V, Bennett MJ, Walker JC, Ma C, Jiang I, Cordon-Cardo C, Li QJ et al (2009) miR-19 is a key oncogenic component of mir-17-92. Genes Dev 23(24):2839–2849

Oshima G, Poli EC, Bolt MJ, Chlenski A, Forde M, Jutzy JMS, Biyani N et al (2019) DNA Methylation Controls Metastasis-Suppressive 14q32-Encoded miRNAs. Cancer Res 79(3): 650–662

Rooj AK, Ricklefs F, Mineo M, Nakano I, Chiocca EA, Bronisz A, Godlewski J (2017) MicroRNA-Mediated dynamic bidirectional shift between the subclasses of glioblastoma Stem-like cells. Cell Rep 19:2026–2032

Roth SA, Knutsen E, Fiskaa T, Utnes P, Bhavsar S, Hald ØH, Løkke C et al (2016) Next generation sequencing of microRNAs from isogenic neuroblastoma cell lines isolated before and after treatment. Cancer Lett 372(1):128–136

Selbach M, Schwanhäusser B, Thierfelder N, Fang Z, Khanin R, Rajewsky N (2008) Widespread changes in protein synthesis induced by microRNAs. Nature 455(7209):58–63

Shahar T, Granit A, Zrihan D, Canello T, Charbit H, Einstein O, Rozovski U et al (2016) Expression level of miRNAs on chromosome 14q32.31 region correlates with tumor aggressiveness and survival of glioblastoma patients. J Neuro-Oncol 130(3):413–422

Soriano A, Masanas M, Boloix A, Masiá N, París-Coderch L, Piskareva O, Jiménez C et al (2019) Functional high-throughput screening reveals miR-323a-5p and miR-342-5p as new tumor-suppressive microRNA for neuroblastoma. Cell Mol Life Sci 76(11):2231–2243

Tsang JS, Ebert MS, van Oudenaarden A (2010) Genome-wide dissection of microRNA functions and cotargeting networks using gene set signatures. Mol Cell 38(1):140–153

Uziel T, Karginov FV, Xie S, Parker JS, Wang YD, Gajjar A, He L et al (2009) The miR-17~92 cluster collaborates with the Sonic Hedgehog pathway in medulloblastoma. Proc Natl Acad Sci U S A 106(8):2812–2817

Valastyan S, Reinhardt F, Benaich N, Calogrias D, Szász AM, Wang ZC, Brock JE, Richardson AL (2009) Weinberg R.A.A pleiotropically acting microRNA, miR-31, inhibits breast cancer metastasis. Cell 137:1032–1046

Van Roosbroeck K, Calin GA (2017) Cancer Hallmarks and MicroRNAs: The Therapeutic Connection Adv. Cancer Res 135:119–149

Weeraratne SD, Amani V, Teider N, Pierre-Francois J, Winter D, Kye MJ et al (2012) Pleiotropic effects of miR-183~96~182 converge to regulate cell survival, proliferation and migration in medulloblastoma. Acta Neuropathol 123(4):539–552

Yang D, Sun Y, Hu L, Zheng H, Ji P, Pecot CV, Zhao Y et al (2013) Integrated analyses identify a master MicroRNA regulatory network for the mesenchymal subtype in serous ovarian cancer. Cancer Cell 23:186–199

Chapter 11
3D Modeling of Non-coding RNA Interactions

Krishna Pal Singh and Shailendra Gupta

Abstract Non-coding RNAs (ncRNAs) are a growing class of transcripts, with lengths ranging from tens to several thousand of bases, involved in the regulation of a large number of biological processes and diseases. Many of these ncRNAs have emerged as the molecules of interest for prognostic, diagnostic, and therapeutic purposes in many diseases including cancer. Although ncRNAs do not encode proteins, they fold into complex structures to interact with target proteins, DNA, or other RNAs. In contrast to microRNAs (miRNAs) where researchers mainly focused on the nucleotide sequence for target prediction in the past, folding and structural conservation seems to be important to encode functions and interactions of long non-coding RNA (lncRNA). In this chapter, we discuss methods and tools available for the structural modeling of ncRNAs together with various examples from the literature where structural modeling helped decipher the function of ncRNAs. We also provide a step-by-step procedure to design 3D structures of ncRNAs combining state-of-the-art tools available toward the design of novel RNA therapeutics.

Keywords Non-coding RNAs · Structure modeling · miRNA-mRNA · Molecular docking · Molecular dynamic simulation · Deep learning

11.1 Introduction

The Human Genome Project revealed that our genome is mostly composed of non-protein-coding DNA. Functional non-coding RNAs (ncRNAs), which are transcribed from DNA, do not encode proteins but play important roles in the regulation

K. P. Singh
Department of Systems Biology and Bioinformatics, University of Rostock, Rostock, Germany

S. Gupta (✉)
Department of Systems Biology and Bioinformatics, University of Rostock, Rostock, Germany

Chhattisgarh Swami Vivekananda Technical University, Bhilai, Chhattisgarh, India

U. Schmitz et al. (eds.), *Systems Biology of MicroRNAs in Cancer*, Advances in Experimental Medicine and Biology 1385,
https://doi.org/10.1007/978-3-031-08356-3_11

of various cellular processes and disease pathophysiology. Together, ncRNA transcripts constitute almost 60% of the transcriptional output in human cells (Feingold et al. 2004; Sarah et al. 2012). There are thousands of ncRNAs that have been identified, which are classified into different categories based on their functions and lengths (Cech and Steitz 2014). These include transfer RNA (tRNA), ribosomal RNA (rRNA), circular RNA (circRNA), small nucleolar RNA (snoRNA), small nuclear RNA (snRNA), microRNA (miRNA), and long ncRNA (lncRNA) (Cheng et al. 2005; Washietl et al. 2007). ncRNAs have been found to play key roles in the regulation of tumor-associated pathologies through different processes and mechanisms. These include transcriptional and posttranscriptional regulation of tumor genes, chromatin remodeling, and signal transduction. Like proteins, RNA also folds into well-defined three-dimensional (3D) structures. To understand their function, to design synthetic RNA-based regulators, and to discover drugs targeting functional RNAs, knowledge of their structure is crucial.

In the last two decades, the functional annotations of RNA molecules were mainly based on their secondary (2D) structure confirmations represented by Watson–Crick (WC) base pairing for which several methods have been previously developed (Zuker 2003; Ding et al. 2008; Mathews et al. 2010). For RNA sequences with <700 nucleotides, the secondary structure of RNA molecules can be determined based on thermodynamic principles with almost 70% accuracy (Mathews 2004). This limited accuracy is attributed to the fact that thermodynamics alone is not the only determinant of free energy change. The main challenge is to determine the optimal structural fold with base pairings resulting in the lowest free energy change from an unfolded to folded state of the RNA molecule. This task becomes increasingly complex due to an exponentially increasing folding space with longer RNA sequences such as lncRNAs, which can range from a few hundreds to several thousand nucleotides in length.

For experimental determination of RNA structures, chemical probing-based methods have been used in the past several decades (Draper et al. 2005; Weeks 2010). Over the years, chemical probing methods have been extended to illustrate the native state and flexibility of RNA molecules inside the living cells in a high-throughput manner (Rouskin et al. 2013; Loughrey et al. 2014; Talkish et al. 2014; Kubota et al. 2015; Lorenz et al. 2016) including their three-dimensional structure assessment (Weeks 2010; Cordero and Das 2015). Still, there is only a little progress made on the 3D structural elucidation of RNAs experimentally. As shown in Fig. 11.1, there is a huge gap between the number of experimentally determined nucleic acid structures available in Protein Data Bank (PDB) in comparison to protein structures. Even though the human non-coding transcriptome is approximately 30 times larger than that coding for proteins, the number of experimentally determined RNA structures is <1% of total protein structures resolved. If we search for ncRNA 3D structures, currently there are only 47 experimentally curated structures available in PDB. These numbers indicate that most common methods for the computational modeling of protein structures such as homology modeling cannot be simply extended in case of ncRNA structures due to the absence of suitable templates.

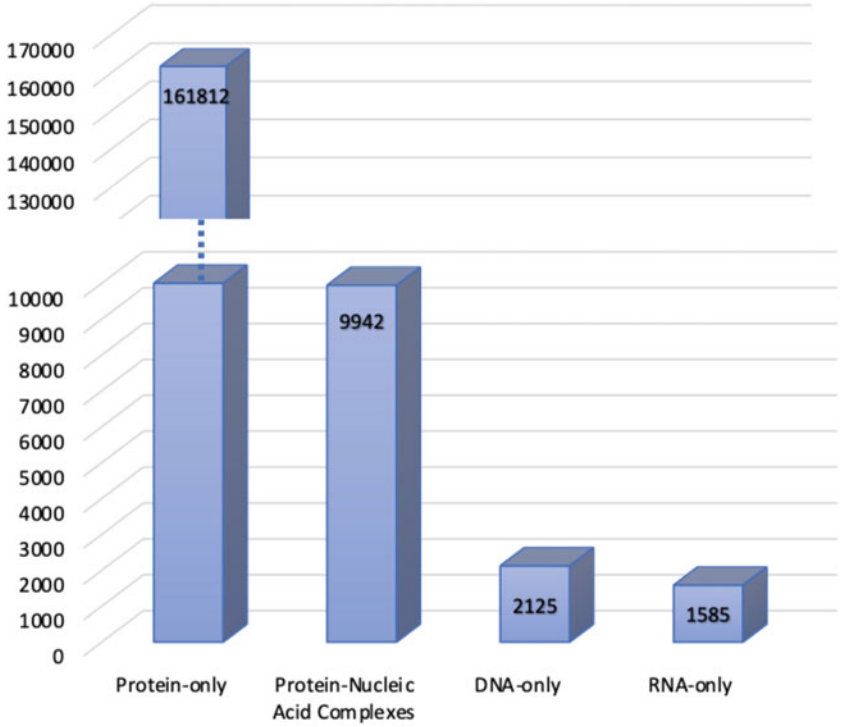


Fig. 11.1 Experimentally determined structures of biomolecules in PDB as of December 2021

11.2 Structure Modeling of ncRNAs

The number of ncRNAs are expanding rapidly, so do the computational tools to predict their structure in the 3D space. Considering the total ncRNA universe known so far, the Rfam repository (http://rfam.xfam.org) has 4069 RNA families across all species each of which is represented by manually curated sequence alignments, consensus secondary structures, and predicted homologues (Kalvari et al. 2018, 2021). Interestingly, in the Rfam database, for members of an RNA family the secondary structure is more conserved than their primary sequence. However, the identification of a functional RNA structure purely based on minimum folding energy (MFE) calculations is generally considered not reliable in the absence of experimental validation.

Computer programs like RNAfold (available via the ViennaRNA Web Services; http://rna.tbi.univie.ac.at) can provide the MFE structure of any given RNA sequence (up to the length of 10,000 bases) using the loop-based energy model and the dynamic programming algorithm. Although, the MFE structure of ncRNAs can be considered somewhat more stable, the difference in the folding energies between the native ncRNA sequence and a random sequence created by reshuffling

of the original sequence is often too indistinguishable with the exception of miRNAs (Workman and Krogh 1999; Rivas and Eddy 2000).

Like other biomolecules, ncRNAs also take a 3D confirmation in nature, which can be predicted from the 2D structure by looking for the connections between various bases in the MFE structure. Considering this, the MFE secondary structure represents a useful abstraction of the full-length 3D structure of RNA molecules (Gorodkin and Hofacker 2011).

Various tools and web services for designing ncRNA structures and predicting interaction sites with proteins, DNA, and other RNAs are provided in Table 11.1.

11.3 Modeling of miRNA–mRNA–Argonaute complexes

Among non-coding RNAs, miRNAs are involved in the regulation of a large number of developmental and physiological processes by sequence-specific recognition and inhibition of target mRNA. The central component of this targeted mRNA repression is the miRNA-induced silencing complex (miRISC) formed between miRNA, Argonaute (Ago), and other Ago binding proteins including GW182. The crystal structure of human Argonaute 1 (Ago1, PDB ID: 4KRF) and Ago2 (PDB ID: 4EI1) are already resolved. To understand the structural basis of miRNA-mediated silencing, Schirle and colleagues have experimentally determined the crystal structures of human Ago2–miRNA complexes (PDB IDs: 4W5N, 4W5O, 4W5Q, 4W5R, and 4W5T) with and without target mRNA (Schirle et al. 2014). These experimental structures provide a foundation to prepare and analyze computational models of various miRNA–mRNA duplexes in complex with Ago for more accurate assessment of miRNA function.

A large number of algorithms have been developed for miRNA-target recognition based on the base pairing in the seed region and/or duplex secondary structure (Quillet et al. 2020). Among them, miRanda (Betel et al. 2010), PITA (Kertesz et al. 2007), SVmicrO (Liu et al. 2010), TargetScan (Agarwal et al. 2015), MBSTAR (Bandyopadhyay et al. 2015), miRWalk (Dweep and Gretz 2015), ExprTarget (Gamazon et al. 2010), and miRMap (Vejnar and Zdobnov 2012) are frequently used by researchers. However, there is increasing experimental evidence for the frequent use of non-canonical binding sites for miRNAs with imperfect seed pairing or even formation of seedless duplexes (Helwak et al. 2013; Chipman and Pasquinelli 2019; Sheu-Gruttadauria et al. 2019). These duplexes can often not be predicted using conventional computational methods, which rely on the sequence features.

Several structure-based analyses indicate that miRISC not only stabilizes miRNAs but also coordinates the formation of miRNA-target base pairing (Wang et al. 2008; Parker et al. 2009) and thereby facilitates the miRNA–mRNA duplex even in cases with imperfect seed binding sites. Combining structural modeling of Ago with miRNA–mRNA duplexes with features from sequence and secondary

Table 11.1 Tools and webservices frequently used for the design and analysis of ncRNA structures and interactions

Name	Description	Link	References
Secondary and tertiary structure prediction tools for ncRNAs			
MINT: Motif identifier for nucleic acid trajectory	Tool for analyzing 3D structures of nucleic acids and their full atom molecular dynamics trajectories. Predicts secondary confirmation of RNA and RNA motifs.	Mint.Cent.Uw.Edu.Pl	Górska et al. (2015)
VARNA: Visualization applet for RNA	Tool for drawing, visualization, and annotation of secondary structures of RNA	varna.lri.fr	Darty et al. (2009)
Forma: Force-directed RNA	RNA secondary structure visualization tool	rna.tbi.univie.ac.at/forna	Kerpedjiev et al. (2015)
ViennaRNA web services	Collection of tools to predict minimum free energy secondary structure of RNA (limit 10 K nt)	rna.tbi.univie.ac.at	Gruber et al. (2015)
NUPACK	Tool for analysis and design of nucleic acid structures (limit 30 K nt)	nupack.org	Zadeh et al. (2011)
3DRNA	RNA and DNA tertiary structure and non-coding RNAs	biophy.hust.edu.cn/new/3dRNA	Zhang et al. (2020)
RNAComposer	Fully automated prediction of large RNA 3D structures	rnacomposer.cs.put.poznan.pl	Biesiada et al. (2016)
SimRNA	RNA 3D structure modeling with optional restraints	genesilico.pl/software/stand-alone/simrna/	Boniecki et al. (2016)
Molecular docking tools for ncRNA/protein interactions			
P3DOCK	Protein-RNA docking based on a hybrid algorithm of template-based (PRIME) and free docking (3dRPC)	rnabinding.com/P3DOCK/P3DOCK.html	Zheng et al. (2020)
HDOCK	Protein–protein and protein–DNA/RNA docking based on a hybrid algorithm of template-based modeling and ab initio free docking	hdock.phys.hust.edu.cn	Yan et al. (2017)
PATCHDOCK and FIREDOCK	Molecular docking algorithm based on object recognition and image segmentation techniques. The server can be used for protein, DNA, RNA, peptide, and drug interactions. The interactions posed are further refined using FIREDOCK. These two programs can be utilized for ncRNA–protein interactions	bioinfo3d.cs.tau.ac.il/PatchDock http://bioinfo3d.cs.tau.ac.il/FireDock	Schneidman-Duhovny et al. (2005), Mashiach et al. (2008)

(continued)

Table 11.1 (continued)

Name	Description	Link	References
HADDOCK	HADDOCK is an information-driven flexible docking program for the modeling of a variety of biomolecular complexes. HADDOCK can be used for the interaction of ncRNAs and proteins	https://wenmr.science.uu.nl/haddock2.4/	Dominguez et al. (2003)
MPRDOCK	MPRDOCK server predicts the complex structure between proteins and nucleic acids by efficient consideration of protein flexibility	huanglab.phys.hust.edu.cn/mprdock/	He et al. (2019)
Hexserver	FFT-based algorithm to exploit shape complementarity as a feature for optimization	hexserver.loria.fr	Macindoe et al. (2010)
AutoDock	For the prediction of a variety of biomolecular complexes. The program can be adopted for protein-ncRNA interactions	https://autodock.scripps.edu/download-autodock4/	Morris et al. (2009)
ZDOCK	ZDOCK optimizes desolvation, shape complementarity, and electrostatics using the FFT-algorithm to identify interactions between two proteins. The program can also be adopted for protein–ncRNA interactions	zdock.umassmed.edu	Pierce et al. (2014)
Sequence level ncRNA/protein interaction predictions			
CatRAPID	Predicts protein interactions with large RNA	s.tartaglialab.com/update_submission/428105/d4ee4f48bd	Yan et al. (2017)
RPISeq	Computational methods for predicting RNA–protein interactions	pridb.gdcb.iastate.edu/RPISeq/	Muppirala et al. (2011)
Force fields are frequently used for ncRNA molecular dynamic simulation			
AMBER	Widely used force fields for ncRNA simulations. AMBER ff94 was initially developed for RNA simulations which were further refined with several dihedral parameters in ff98 and ff99. The recent force field used for RNA is ff99bsc0xOL3. New	ambermd.org/AmberModels.php	Cheatham et al. (1999), Wang et al. (2000), Bayly et al. (2002), Šponer et al. (2017)

(continued)

Table 11.1 (continued)

Name	Description	Link	References
	parameters are aimed mainly at refining specific torsion angles and certain non-bonded terms		
CHARMM	The CHARMM36 force field parameters are widely used to simulate nucleic acids, proteins, and other heterogeneous biomolecular system. CHARMM force fields are commercialized and available in the Biovia Discovery Studio software suit (https://www.3ds.com)	mackerell.umaryland.edu/charmm_ff.shtml	Xu et al. (2016), Huang et al. (2017)

structure levels might help in predicting non-canonical and seedless duplexes more accurately.

Gan and Gunsalus used structure modeling methods to investigate the role of Ago in the formation of miRISC (Gan and Gunsalus 2015). Using human and *C. elegans* Ago structures, they showed that Ago adopts variable conformations to generate distorted and imperfect miRNA–target duplexes. For the analysis, they first generated open Ago conformation using the *elNemo* implementation of the Elastic Network Model (ENM), which is available at http://www.sciences.univ-nantes.fr/elnemo/ and selected a structure from low-frequency mode to accommodate a miRNA–mRNA duplex structure. They generated RNA duplex structures using the MC-Sym algorithm (Parisien and Major 2008). Further, for each of the duplex structures, 3D structure models were prepared and ranked using an all-atom AMBER99 force field with atomic interaction and implicit solvation energies. Further, they loaded the duplex into the open Ago structures by superimposing a guide RNA strand to the experimentally resolved *T. thermophilus* Ago-duplex structure (PDB ID: 3HJF) as a reference structure using the superpose routine of the TINKER package (https://dasher.wustl.edu/tinker/) (Pappu et al. 1998). Finally, they superimposed human and *C. elegans* Ago models to the reference structure using combinatorial extension (CE) algorithms (Shindyalov and Bourne 2001) for preparing final miRISC structures. A computational pipeline used by Gan and Gunsalus for the tertiary structure analysis of miRNA–target interactions is shown in Fig. 11.2.

The human and *C. elegans* miRISC structures modeled by Gan and Gunsalus using experimentally a solved bacterial miRISC structure suggest that both human and *C. elegans* Ago proteins have similar dynamic modes of motion compared to bacterial Ago. Interestingly, they find similar binding affinity treads for loaded RNA duplexes for *T. thermophilus*, yeast, and human miRISCs. Their miRISC models also highlight that the quality of duplex accommodation within Ago structures

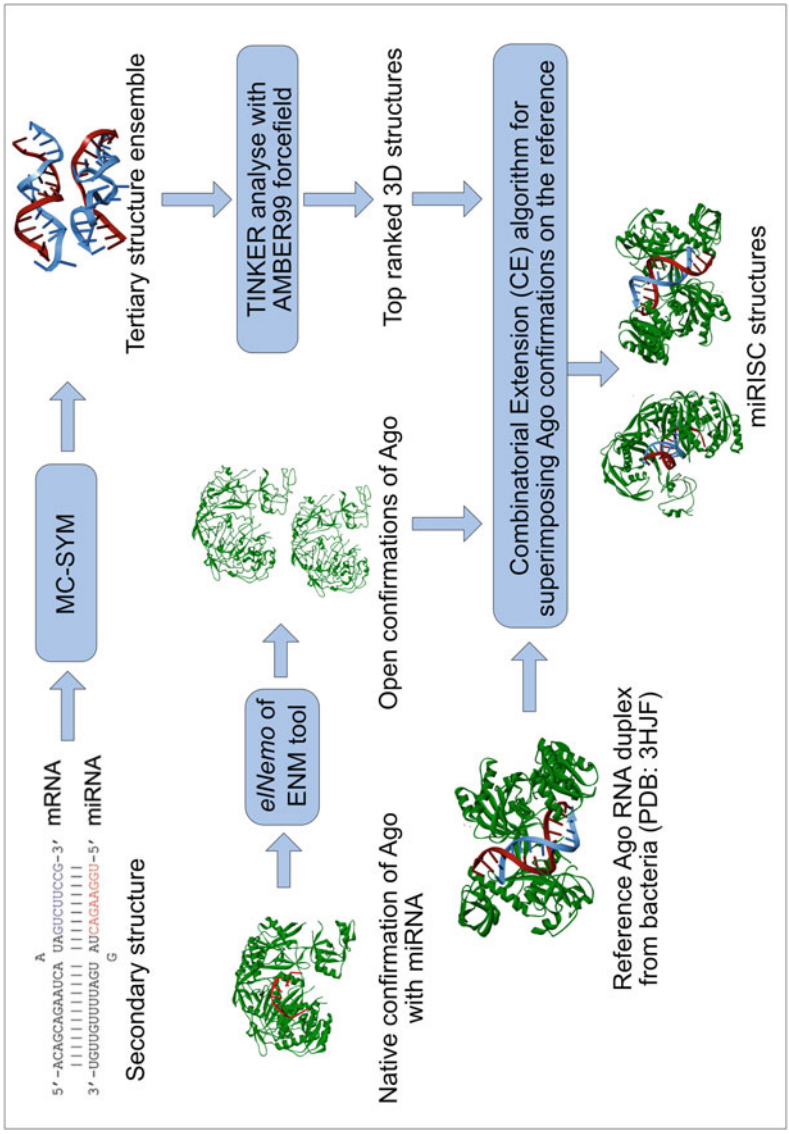


Fig. 11.2 Computational pipeline used for predicting the new miRISC structures. Tools used in the computational pipeline are shown in the blue boxes. The figure is redesigned from Gan and Gunsalus (2013, 2015)

depends on the duplex structure distortions. Structural distortions in the seed region (2–8) and 3′ end regions (nt positions 14–21) cause less steric clashes with Ago compared to the central region (nt position 10–14). In their computational pipeline, the miRNA–mRNA duplexes and Ago were superimposed on the already available reference bacterial Ago–miRNA–mRNA complex. However, considering various sequence lengths and arrangements of bases both in miRNAs and their target sites, simple superimposition of the duplexes on the already resolved miRISC structure might not be a universal solution to design and predict new miRISC structures.

11.4 Computational Pipeline for the Structural Modeling of ncRNAs, Proteins, and Their Interactions

In this section, we provide a step-by-step procedure to design 3D structures of ncRNA and proteins with open access software and web-based services followed by prediction of their interaction interface using molecular docking and molecular dynamics simulations.

1. *Extraction of ncRNA and protein sequences*

   For modeling ncRNA and protein 3D structures and interactions, the first step is to retrieve the sequence of the molecules of interest from public databases (Table 11.2).

Table 11.2 Protein and ncRNA sequence databases

Database	Brief description	Link
NCBI Refseq	Comprehensive, integrated, non-redundant, well-annotated set of sequences, including genomic DNA, transcripts, and proteins.	www.ncbi.nlm.nih.gov/refseq/
UniProt	Comprehensive resource for protein sequence and annotation data. The UniProt databases are the UniProt Knowledgebase (UniProtKB), the UniProt Reference Clusters (UniRef), and the UniProt Archive (UniParc).	www.uniprot.org/
Protein data Bank	Archive information about the 3D shapes of proteins, nucleic acids, and complex assemblies.	www.rcsb.org/
miRBase	Published archive of microRNA sequences and annotations.	www.mirbase.org/
NONCODE	Contains a total of 487,164 lncRNA transcripts and 324,646 lncRNA genes for 39 different species.	www.noncode.org/
LNCipedia	LncRNA sequence and annotation database. The current release contains 127,802 human lncRNAs transcripts.	lncipedia.org/
RNAcentral	Comprehensive annotations of eukaryotic lncRNAs.	rnacentral.org
GENCODE	Contains ncRNA gene annotations in GTF format and ncRNA transcript sequences in FASTA format.	www.gencodegenes.org/

For ncRNAs, RNAcentral is one of the most widely used public resources that offers integrated access to a comprehensive and up-to-date set of non-coding RNA sequences provided by a collaborating group of Expert Databases representing a broad range of organisms and RNA types (Fig. 11.3) (Petrov et al. 2015). Most of these databases offer the utility to download ncRNA and protein sequences in the FASTA format suitable for the secondary structure prediction and structure modeling tools highlighted in Table 11.1.

2. *Structure modeling of ncRNAs*

   Currently, most of the state-of-the-art tools to design the 3D structure of ncRNAs are based on the thermodynamically stable secondary structure of the RNA sequence. For this purpose, the RNAfold server, available within the ViennaRNA web services, is frequently used by researchers among many other tools as highlighted in Table 11.1. After predicting the thermodynamically stable secondary structure of an ncRNA, the next step is to predict the ncRNA tertiary structure. MC-Fold/MC-Sym, 3DRNA, and RNAComposer are the most widely used web-based tools for this purpose. In most cases, the input to the server is simply the RNA sequence in FASTA format followed by its secondary structure in dot-bracket notation. While MC-Fold/MC-Sym has a limit of 150 nucleotides, RNAComposer can handle RNA sequences of up to 500 nt for their 3D structure prediction. Larger ncRNA structures can be modeled using the 3dRNA tool (Fig. 11.4) (http://biophy.hust.edu.cn/new/3dRNA), where the RNA sequence is first decomposed into smallest secondary elements, including helix, hairpin loop, internal loop, bulge loop, pseudoknot loop, and multibranch loop (also known as a junction). The 3D structure of each element is then computed using the appropriate 3D template, which is often experimentally determined using X-ray crystallography or NMR. For those segments where the appropriate 3D template is missing, the structure is prepared ab initio using a distance geometry-based loop building method. Finally, the 3D elements are assembled into an integrated 3D structure based on the information from the 2D structure, followed by minimization of final predicted structure using AMBER force fields in the generalized Born solvent model, to remove any atomic clashes in the model.

3. *Structure modeling of proteins*

   The protein 3D structure for interaction analysis can be extracted directly from the Protein Data Bank (https://www.rcsb.org/). In case the protein 3D structure is not available, a variety of methods are available including homology modeling, threading-based approaches, ab initio, machine learning, or deep learning-based structure prediction. Detailed descriptions of each of the protein 3D structure prediction method are out of the scope of this chapter. There are many detailed review articles that can be explored for the 3D structure prediction of proteins (Dorn et al. 2014; Kuhlman and Bradley 2019; Senior et al. 2020). One of the frequently used web-based servers to predict the protein structures which are not resolved yet is I-TASSER (Fig. 11.5) (zhanggroup.org/I-TASSER/).

   In the recent Critical Assessment of Techniques for Protein Structure Prediction (CASP14) challenge for the assessment of methods of protein structure modeling, the artificial intelligence-based AlphaFold program developed by

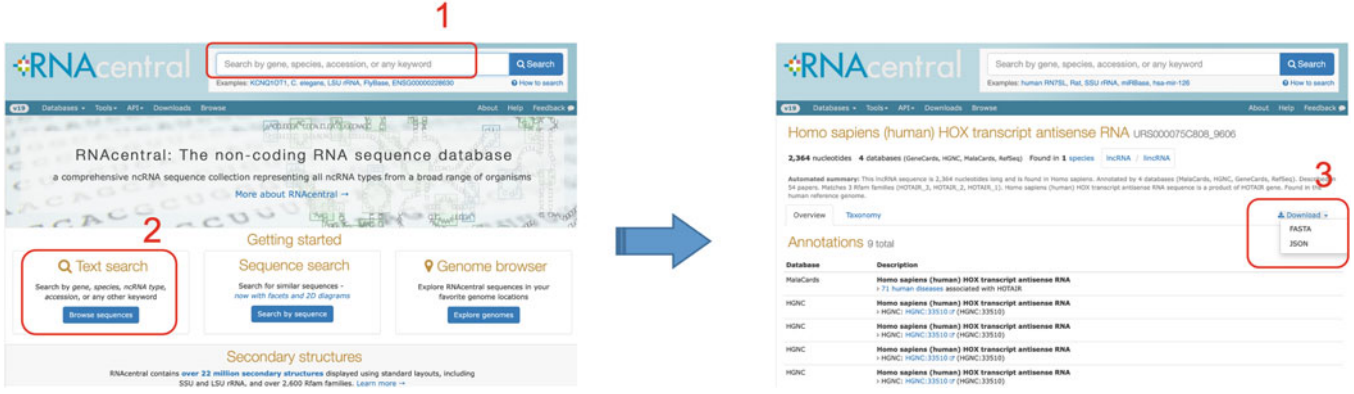


Fig. 11.3 Snapshot of RNAcentral database. ncRNAs can be searched directly by providing the gene name, accession number, or a keyword in the search bar (1). Alternatively, the text search (2) can be used for more advanced search options and search filters. In the example shown in the right panel, we searched the human HOTAIR lncRNA. The sequence can be downloaded directly in FASTA format (3) for further processing

Fig. 11.4 Snapshot of the 3dRNA Web Server. The server needs only two basic inputs from the user for RNA 3D structure modeling (left). These are (1) the sequence of RNA and (2) the secondary structure in dot-bracket notation. For the latter, the server also offers secondary structure prediction from one of the established methods (RNAfold, MaxExpect, Probknot, or IPknot). In the right panel of the figure, a sample output is shown. The 3dRNA server computes the top five minimized models (by default), which can be downloaded in the PDB format (3)

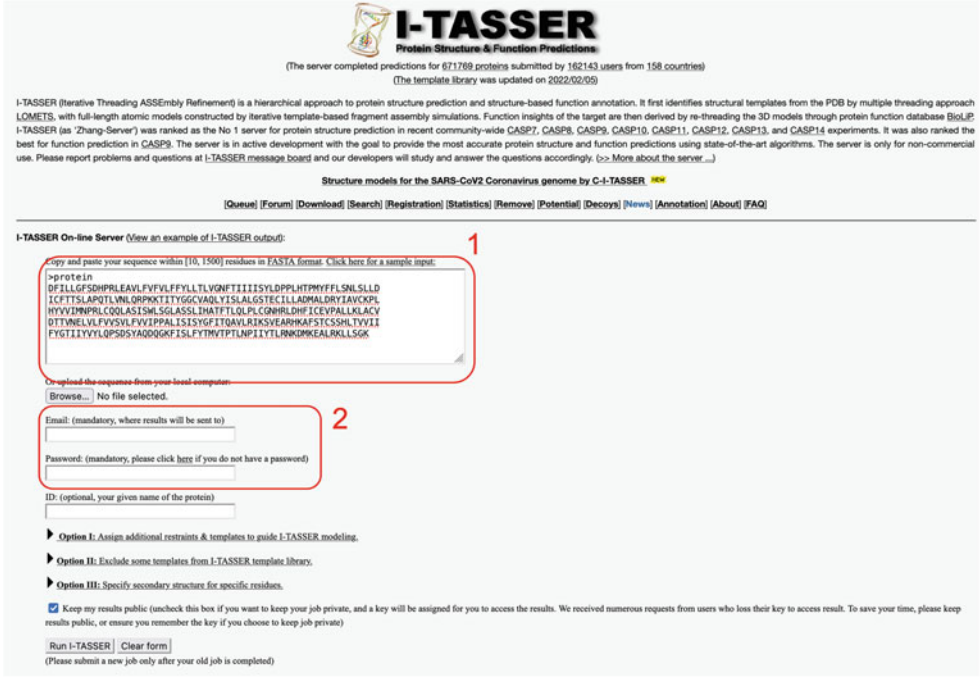


Fig. 11.5 Snapshot of the I-TASSER Server which predicts protein 3D structures. Registered users need to provide the protein sequence in FASTA format (1) along with their credentials (2). There are several advanced parameters that can be fine-tuned for the prediction of protein structures including contact/distance restrains, inclusion/exclusion of some specific templates, setting secondary structure profiles for specific residues, etc. The server computes the top 5 models for each of the protein sequences submitted which are ranked based on the C-score (confidence score for estimating the quality of predicted model by I-TASSER)

Alphabets's/Google's DeepMind (Senior et al. 2020) ranked number one in the protein structure prediction from the amino acid sequence with high accuracy. Other neural network-based protein structure prediction program RoseTTaFold can accurately model multichain protein–protein complexes (Baek et al. 2021).

4. *Prediction of protein and non-coding RNA binding from sequence data*

   For shorter ncRNAs, such as miRNAs, it is possible to perform molecular docking studies directly after the 3D structure of the ncRNA and protein are prepared. However, for lncRNAs, it is suggested to first identify potential lncRNA–protein interaction sites at sequence level to reduce the search space in molecular docking programs. There are several tools available (Table 11.1) that, based on sequence information, predict potential ncRNA/protein binding sites. Some widely used tools for lncRNA–protein interaction site predictions include CatRAPID (Armaos et al. 2021), RPISeq (Muppirala et al. 2011), lncPro (Lu et al. 2013), RPI-Pred (Suresh et al. 2015), and rpiCOOL (Akbaripour-Elahabad et al. 2016). CatRAPID estimates the binding propensity of protein–RNA pairs by combining secondary structure, hydrogen bonding, and van der Waals contributions. RPISeq and RPI-Pred predictions are based on Support Vector Machine (SVM), while rpiCOOL predicts binding regions based on sequence motifs. Putative binding sites between lncRNA and proteins can be further prioritized based on features of lncRNA and protein secondary and tertiary structures, such as surface charge and solvent accessibility. Predicted sites can be used as input in molecular docking programs to fine-tune lncRNA–protein interaction poses.

5. *Molecular docking of ncRNAs and proteins*

   Molecular docking tools play an important role in the prediction and investigation of binding interfaces and key residues involved in the interaction of molecules. For the molecular docking of biomolecules, different conventional tools and webservers are available which can be used for ncRNA–protein interaction analysis (Table 11.1).

   Web-based servers such as Hex (hex.loria.fr/), which was initially designed for protein–DNA and protein–ligand docking based on 3D shape similarity and complementarity calculation, are now used for protein–ncRNA interactions as well, considering ncRNAs as a rigid molecule (Wang et al. 2010; Feng et al. 2019; Philip et al. 2021). Conventional molecular docking tools, such as Autodock (autodock.scripps.edu) are already being adopted to identify interaction sites between proteins and miRNAs (Wang et al. 2010; Suyal et al. 2018; Mallick et al. 2019; Mahernia et al. 2020) and lncRNA–ligand interactions (Abulwerdi et al. 2019; Li et al. 2019). Like Hexserver, ZDOCK, which uses a Fast Fourier Transform algorithm to enable global docking search on a 3D grid in combination with shape complementary and electrostatics parameters for protein–protein docking (Pierce et al. 2014) was also successfully employed for protein–ncRNA interactions (Bose et al. 2015; Iwakiri et al. 2016; Jha et al. 2020; Koralewska et al. 2021).

   Programs such as HADDock (wenmr.science.uu.nl/haddock2.4/) allow flexible docking incorporating knowledge from experimental and computational data

to drive the modeling process. Users can provide hints about the interaction regions from, for example, mutagenesis, mass spectrometry, NMR, chemical shift perturbation, or from various computational analyses such as prediction of protein–RNA binding sites and structural motifs. HADDOCK uses these features as ambiguous interaction restraints to identify interactions between protein–protein, protein–nucleic acid, and with small molecules. Jiang and colleagues used HADDOCK to identify initial binding between human Ago and miRNAs (Jiang et al. 2015). Salerno and colleagues exploited Hex and HADDOCK to show that the lncRNA DLEU2 interacts with the Hepatitis B protein HBx and the histone methyltransferase Enhancer of Zeste Homolog 2 (EZH2) (Salerno et al. 2020). Eichhorn and colleagues used HDOCK to explore interactions between human La-related protein group 7 (hLarp7) and 7SK lncRNA (Eichhorn et al. 2018).

The HDOCK program (hdock.phys.hust.edu.cn) uses hybrid docking strategies based on template-based as well as ab initio docking to predict the binding complexes between two molecules including proteins and nucleic acids. The server is capable of both, blind and controlled docking between two molecules depending on the availability of binding site information. Naderi et al. used HDOCK to explore the interactions between two lncRNAs (NONHSAT139215 and NONHSAT139219), which were significantly downregulated in patients with severe hemophilia A, with coagulation factor VIII (Naderi et al. 2018). Similarly, Yeh et al. used HDOCK to understand the regulatory role of various fragments of lncRNA NDRG1-OT1, which is upregulated under hypoxia, on NDRG1 promoter and other NDRG1 regulating proteins including HNRNPA1, KHSRP, and HIF-1α (Yeh et al. 2018). Recently, Lu et al. used HDOCK to explore how single nucleotide polymorphisms and variants alter the binding affinities of lncRNA HCG23 and transcription factor E2F6 (Lu et al. 2020).

Other frequently used molecular docking tools that are also employed for protein–miRNA and protein–lncRNA interactions are PatchDock (https://bioinfo3d.cs.tau.ac.il/PatchDock/) and FireDock (http://bioinfo3d.cs.tau.ac.il/FireDock/). Ghosh and colleagues used PatchDock to model the interaction of NEAT1 lncRNA with the JunD transcription factor and ZO-1 promoter (Ghosh et al. 2021). Similarly, Bozgeyik and colleagues performed interaction studies between miR-19a-3p and miR-421 with Ago protein and PCA3 lncRNA using the PatchDock server (Bozgeyik et al. 2021). In one of our previous studies, we used the PatchDock and FireDock servers to investigate the interactions between SLC16A1-AS1 lncRNA with E2F1 transcription factor on the MCT1 promoter (Logotheti et al. 2020), which is described in detail in the case study section of this chapter.

Besides open access webservers and tools available for molecular docking, commercial docking software such as Schrödinger's Small Molecular Discovery Suite (https://www.schrodinger.com/platform/drug-discovery) for structure preparation and virtual library screening and Dassault Systèmes's Biovia Discovery Studio (https://www.3dsbiovia.com/) are also used to explore interactions between ncRNA–ncRNA; ncRNA–protein, and other biomolecules (François-Moutal et al. 2021).

6. *Molecular dynamic simulations*

The stability of molecular interactions can be assessed using molecular dynamic (MD) simulations, which mimic the physical, chemical, and thermodynamic conditions to predict interactions at atomic level and their dynamics over time using fundamental laws of Newtonian physics (Adcock and McCammon 2006). A comprehensive overview highlighting fundamental methodological challenges and development in the field of RNA MD simulations was recently provided by Šponer and colleagues (Sponer et al. 2018).

The application of conventional and enhanced MD methods on the assessment of ncRNA functions is reviewed in detail by Palermo and colleagues (Palermo et al. 2019). In case of the RNA-guided CRISPR–Cas9 complex, they investigated the mechanistic basis by which the Cas9 protein undergoes structural transitions from an "open" state to a "close" state for RNA binding followed by binding and cleavage of the targeted DNA. They suggested the use of enhanced sampling simulation using the Gaussian accelerated MD (GaMD) method (Miao et al. 2015) that can overcome the time scale limit of MD simulations to capture large-amplitude confirmational changes in case of protein–RNA binding (Palermo et al. 2018).

Other than the time scale limit of traditional simulation engines, another limiting factor in the MD simulation of ncRNA molecules is the availability of molecular mechanical force fields that define the relationship between the molecule's geometry at atomic level and its potential energy. Different classes of force fields have been developed in the past, focusing on different MD simulation systems and types of biomolecules modeled. Some classical force fields include OPLS (Jorgensen et al. 1996), AMBER (Duan et al. 2003), GROMOS (Christen et al. 2005), and CHARMM (MacKerell et al. 1998), which were initially prepared to simulate proteins and peptides and subsequently modified to include nucleic acids as well. AMBER ff94 (Cornell et al. 1995), followed by the correction of several dihedral parameters in AMBER ff98 (Cheatham et al. 1999), and AMBER ff99 (Wang et al. 2000)) are most widely used AMBER nucleic acid force fields. Another class of force field that is used for simulations of ncRNA systems belongs to CHARMM. The latest version of CHARMM is CHARMM36 (Huang et al. 2017), which is tested for nucleic acid simulations. There are several open access and commercial tools that include AMBER and CHARMM force fields to simulate ncRNAs in complex with proteins and other biomolecules. The AMBER force field can be used directly in the AmberTools21 software (https://ambermd.org/) available under the GNU General Public License (GPL). The most widely used open access tool for MD simulations is GROMACS (https://www.gromacs.org/) mainly supporting the simulation of proteins, lipids, and nucleic acids (Abraham et al. 2015). Both AMBER and CHARMM force fields can be used with GROMACS for simulating ncRNA molecules. Another open access, GPU accelerated MD simulation program is NAMD (https://www.ks.uiuc.edu/Research/namd/) developed by the Theoretical and Computational Biophysics Group at the NIH Center for Macromolecular Modeling and Bioinformatics, University of Illinois. NAMD also supports both

AMBER and CHARMM force fields for MD simulation. The Gaussian Accelerated Molecular Dynamics (GaMD) method for unconstrained enhanced sampling and free energy calculations of biomolecules is also supported by both AmberTool21 and NAMD for large-scale MD simulation studies of ncRNA complexes. OpenMM (https://openmm.org/) is another open access program for simulating RNA and protein complexes. In the commercial domain, both Schrödinger's Small Molecular Discovery Suite and Dassault Systèmes's Biovia Discovery Studio provide easy-to-use graphical user interfaces for preparing ncRNA complexes suitable for MD simulations and follow-up analysis. While Schrödinger's Small Molecular Discovery Suite uses the OPLS4 force field (Lu et al. 2021) for RNA and Biovia Discovery Studio contains the commercial version of the CHARMM36 force field (Huang et al. 2017).

11.5 Advances in Modeling Protein–ncRNA Interactions Using Deep Learning

To explore the mechanism of interaction between proteins and ncRNAs, most of the computational pipelines rely heavily on RNA sequence data and thermodynamically stable secondary structure confirmation of RNA molecules along with the use of conventional molecular docking and MD simulation tools. Recently, artificial intelligence-based systems, such as AlphaFold have transformed the entire structural biology domain and open a new gateway to predict 3D structure-level interactions of many unresolved proteins and ncRNAs. Now all the human proteins are modeled using AlphaFold and are currently available in the UniProt database, although many of them are of low quality. However, the prediction capability of deep learning methods will potentially increase in the next few years with the availability of more experimental structures.

Similar to the prediction of protein structures, the newly developed deep learning-based method Atomic Rotationally Equivariant Scorer (ARES) (Townshend et al. 2021) has significantly improved the prediction of RNA structures. After 3D modeling the main task is to determine the molecular function, that is, how a specific molecule can interact with another. Deep learning methods were previously used to optimize the CRISPR guide RNA design (Chuai et al. 2018) and play an important role in designing new antimicrobial peptides (Das et al. 2021). In most of the conventional protein–RNA interaction prediction methods, binding regions between interacting protein–RNA molecules are predicted usually using sequence and structural binding motifs either on protein or RNA. Deep learning methods identify the pattern on both protein and RNA simultaneously to predict the most favorable protein–RNA interaction sites (Lam et al. 2019). While the binding site prediction refers to the RNA binding sites on the protein surface at the structural level, the binding preference means to predict the protein binding preference against given RNA sequences.

Table 11.3 List of selected neural network-based tools to predict RNA binding sites in protein sequences and structures

Name of the tool	Prediction type	Model	Feature	References
aPRBind	Binding site	CNN	PSSM and feature vector	Liu et al. (2021)
Graphbind	Binding site	GNN	Graph, feature vector	Xia et al. (2021)
Deepclip	Binding Preference	CNN + BiLSTM	One-hot encoding	Bjørnholt Grønning et al. (2020)
DeepBind	Binding Preference	CNN	PWM	Alipanahi et al. (2015)
DeeperBind	Binding Preference	CNN	Long short termmemory networks (LSTMs)	Hassanzadeh and Wang (2016)
iDeepS	Binding Preference	CNN	DBN-kmer	Pan et al. (2018)
iDeepE	Binding Preference	CNN	PWM	Pan and Shen (2018)

Overall, methods for modeling molecular interactions can be divided into two categories: The first category is based on the principle that similar structures may have similar functions, which underlies the template-based method to predict binding sites (Yang et al. 2013; Chen et al. 2014; Wu et al. 2018; Xie et al. 2020) and binding preference (Zheng et al. 2016). Performance of these methods depends heavily on the availability of homologous sequences (Senior et al. 2020). The second category of methods combines hand-crafted features with shallow-learning methods, such as Support Vector Machine (SVM) (Maticzka et al. 2014; Zhang et al. 2017; Jolma et al. 2020), logistic regression (Hiller et al. 2006; Kazan et al. 2010; Orenstein et al. 2016; Yan and Kurgan 2017), and random forest (Sun et al. 2016; Li et al. 2017) to explore the binding sites between proteins and RNAs.

To identify the binding preference of proteins for RNAs, several tools have been developed based on Convolutional Neural Networks (CNNs) to detect RNA binding motifs in protein sequences as shown in Table 11.3. For example, the DeepBind tool (Alipanahi et al. 2015) is based on CNN which takes RNA sequences as input and predicts relevant RNA binding proteins. Another tool, DeeperBind (Alipanahi et al. 2015), is based on the Long Short-Term Memory (LSTM) that includes layers into the DeepBind architecture, which is helpful in understanding the long-range dependency between the sequence features extracted by the CNN layers. Similarly, the iDeepS (Pan et al. 2018) tool combines CNN and a bidirectional LSTM to simultaneously identify protein binding sites and structure motifs on RNA sequences.

For RNA binding site prediction on protein surfaces, features such as sequence one-hot encodings (Yan et al. 2016), Position-Specific Scoring Matrix (PSSM), and conservation entropy derived from PSSM are used. Similarly, in case of RNA, RNA feature extraction, much related to the proteins, using the one-hot encodings k-mer models (Orenstein et al. 2016), and PWM (position weight matrix)(Kazan et al.

2010; Orenstein et al. 2016) are used by various tools. In a recent development by Liu et al., a convolutional neural network-based ab initio method for RNA binding residue prediction on proteins is trained by combining features from sequences and structures predicted by I-TASSER (Liu et al. 2021). Xia et al. developed another tool (GraphBind) based on hierarchical graph neural networks for recognizing nucleic acid-binding residues on proteins (Xia et al. 2021).

11.6 Case Study

In this section, we present methodologies from our previously published work to guide researchers in setting up new ncRNA–ncRNA and ncRNA–protein interaction studies.

11.6.1 SLC16A1-AS1 lncRNA Interacts with Transcription Factor E2F1 and Modulates Its Activities

LncRNAs are acknowledged for their role in the regulation of genes, proteins, and miRNAs. In many of our experimental studies, we observed a positive correlation between expression profiles of SLC16A1-AS1 lncRNA with the E2F1 regulated genes involved in cancer metabolism (Logotheti et al. 2020). To support the hypothesis that SLC16A1-AS1 has the potential to interact with E2F1 and facilitate its binding to the promotor region of target genes; we used a computational pipeline that incorporates sequence and structural level analysis of this lncRNA and its interaction with E2F1.

11.6.1.1 Retrieval of Sequence and Secondary Structure Prediction of SLC16A1-AS1

The SLC16A1-AS1 sequence was retrieved from the National Center for Biotechnology Information (NCBI) database (Refseq ID: NR_103743). The SLC16A1-AS1 transcript sequence contains 1521 bases which we submitted to the RNAfold server available on the ViennaRNA web services platform (Gruber et al. 2008; Lorenz et al. 2011) for the identification of the thermodynamically most stable secondary structure (Fig. 11.6).

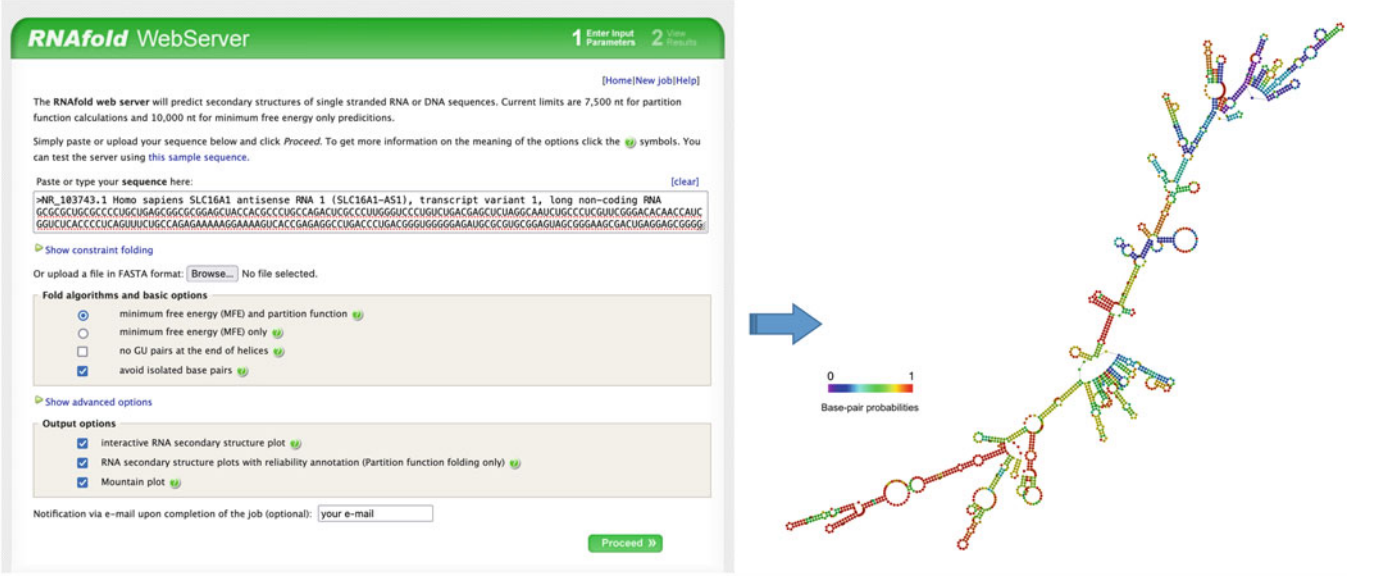


Fig. 11.6 RNAFold webserver is used for the prediction of the thermodynamically most stable secondary structure of SLC16A1-AS1. The structure is shown in the right panel where the nucleotides are colored based on base pair probability

11.6.1.2 Preparation and Optimization of the SLC16A1-AS1 and E2F1 Tertiary Structure

The interaction between the biological molecules and their functional implications can be better understood at the tertiary structure level. For this, we first generated the full-length 3D structure of the SLC16A1-AS1 lncRNA. The basic 3D structure of lncRNAs is generated using the 3DRNA software tool (Wang et al. 2019), which is based on a fragment assembly approach to build RNA 3D structures utilizing the information from the thermodynamically most stable secondary structure predicted by RNAfold.

The structure was manually curated for missing interactions and bond length after applying the CHARMm force field (Vanommeslaeghe et al. 2010) in an iterative manner using the Biovia Discovery Studio 2017 (DS2017) software suite. The curated structure was further optimized using the "Smart Minimizer protocol" available in DS2017 for a maximum run of 20,000 steps with the "Minimization RMS Gradient Tolerance" of 0.1 kcal / (mol x Å) to exit the minimization routine in case the average gradient is less than or equal to the set cutoff (Fig. 11.7). The tertiary structure of E2F1 was used from our previously published works (Meier et al. 2014; Wang et al. 2016; Goody et al. 2019) which was prepared using the I-TASSER webserver and optimized in the Biovia Discovery Studio software suite.

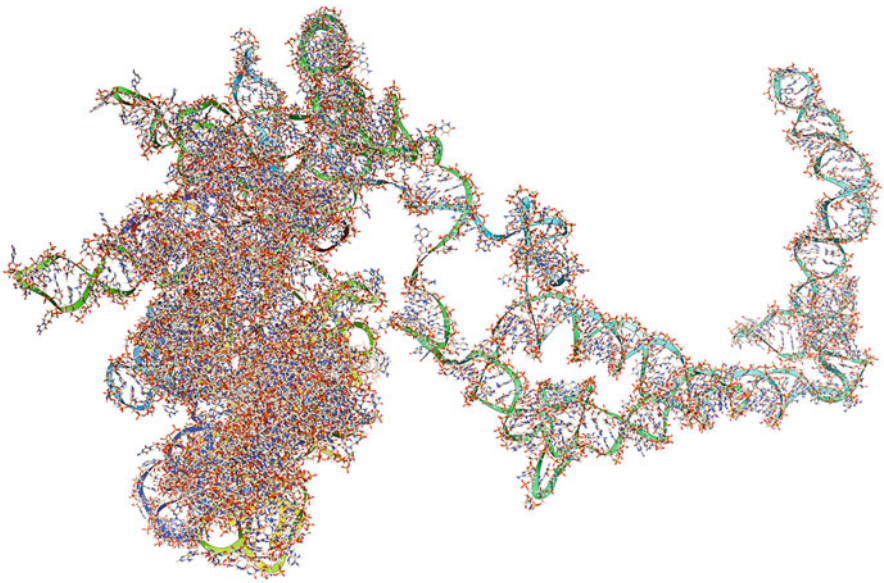


Fig. 11.7 Tertiary structure of full-length SLC16A1-AS1. The model was initially prepared using the 3DRNA software tool and then optimized in Biovia Discovery Studio software using the CHARMm force field. The 3D model is shown as all atom with a flat ribbon backbone

11.6.1.3 Prediction and Prioritization of E2F1 Binding Sites on SLC16A1-AS1 lncRNA

For lncRNAs with large binding surface available for investigating molecular interactions with protein and other biomolecules, it is always advisable to prioritize regions for protein binding to reduce the search space in docking protocols. For the prediction of possible binding sites between E2F1 and SLC16A1-AS1, we used the catRAPID fragment module available on the catRAPID omics server (Agostini et al. 2013). The algorithm was previously tested for the prediction of lncRNA and protein interactions (Bellucci et al. 2011) as shown in Fig. 11.8. The tool first divides protein and RNA into small sequences and then predicts the binding propensity between them in an iterative manner.

For the prioritization of lncRNA regions bound to E2F1, we also considered parameters from the secondary structures of protein and RNA. In case of the lncRNA, loop regions were given priority over the nucleotide forming the stem while in case of the protein, we considered the solvent accessibility parameter calculated using the NetSurfP-2.0 tool (Klausen et al. 2019). The analysis suggested that in the lncRNA fragment (172–233), most of the bases lie in the loop region, while the E2F1 fragment (312–363) has maximum solvent accessibility (absolute solvent accessibility (ASA) = 92.69) among all the possible interacting fragments (Table 11.4).

Based on these clues from the secondary structure of the lncRNA and the solvent accessibility of the protein, lncRNA region between 172–233 nucleotides and E2F1 region between 312 and 363 amino acid residues may be the most promising regions to investigate in a molecular docking analysis.

11.6.1.4 Molecular Docking Between SLC16A1-AS1 and E2F1

After the lncRNA structural modeling and optimization and the identification of a region of interest that has the highest potential to interact with E2F1, we extracted the fragment from 162 to 243 nucleotides from the full-length 3D structure. We explicitly included 10 additional bases at both ends to preserve the folding of the lncRNA fragment. This small fragment was extracted for molecular docking due to the limitation of molecular docking software dealing with protein and nucleic acid interactions. Molecular docking of E2F1 and lncRNA fragments was performed using the PatchDock tool (Schneidman-Duhovny et al. 2005). The top 50 poses identified from the PatchDock tool were refined with the FireDock server (Mashiach et al. 2008). The best pose (Fig. 11.9) based on most favoring thermodynamics of the E2F1 and lncRNA interaction at 25 °C was interestingly in agreement with the binding sites predicted by the catRAPID server on the E2F1 surface using sequence-based information.

In silico mutagenesis experiments for further validation of binding sites between *SLC16A1-AS1* and E2F1.

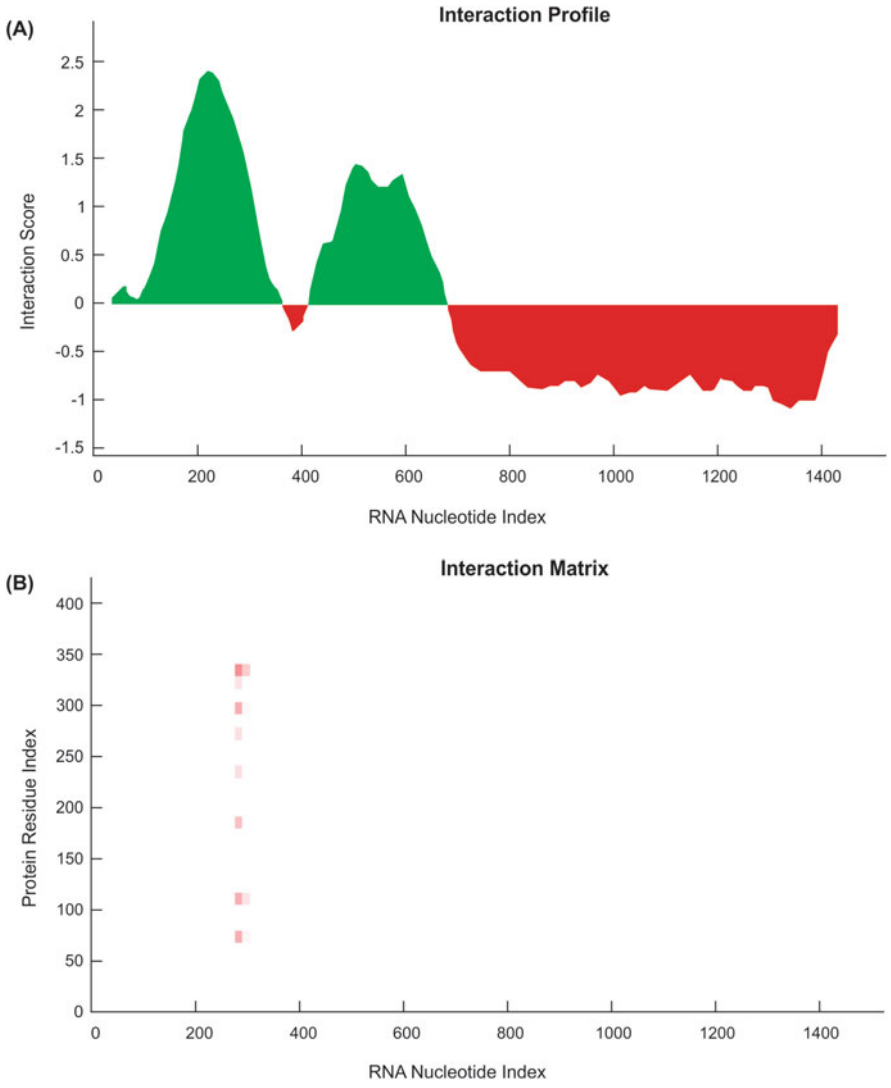


Fig. 11.8 Interaction profile and interaction matrix between SLC16A1-AS1 and E2F1. (**a**) The interaction profile is calculated as an average over protein binding fragments. (**b**) All the protein and lncRNA interacting fragments are shown in interaction matrix

To further support the binding regions identified between E2F1 and lncRNA, we performed in silico mutagenesis experiments by deleting several segments of lncRNA and identify the impact of deletion on E2F1 binding. More specifically, we deleted bases from (1) 172–233; (2) 262–323; (3) 172–233 and 262–323; (4) 193–209; (5) 172–175; and (6) 172–175 and 193–209 together. While the first three deletion studies were associated with the potential binding sites between E2F1

Table 11.4 Top 20 potential interaction sites between E2F1 and SLC16A1-AS1 lncRNA predicted by catRAPID tool

Protein region (E2F1)	RNA region (SLC16A1-AS1)	Interaction propensity	Discriminative power[a]	Absolute surface accessibility of protein region	No. of lncRNA bases forming stem	No. of lncRNA bases forming loop
312–363	241–302	24.17	0.63	92.69209615	45 (72.6%)	17 (27.4%)
51–102	241–302	22.82	0.59	86.88392308	45 (72.6%)	17 (27.4%)
87–138	241–302	22.82	0.59	86.16046154	45 (72.6%)	17 (27.4%)
162–213	241–302	22.02	0.59	60.44236538	45 (72.6%)	17 (27.4%)
276–327	241–302	22.92	0.59	72.46588462	45 (72.6%)	17 (27.4%)
312–363	262–323	21.65	0.57	92.69209615	48 (77.4%)	14 (22.6%)
87–138	262–323	20.48	0.54	86.16046154	48 (77.4%)	14 (22.6%)
212–263	241–302	20.77	0.54	62.41503846	45 (72.6%)	17 (27.4%)
251–302	241–302	20.7	0.54	47.28023077	45 (72.6%)	17 (27.4%)
301–352	241–302	20.57	0.54	90.76119231	45 (72.6%)	17 (27.4%)
51–102	262–323	19.88	0.52	86.88392308	48 (77.4%)	14 (22.6%)
276–327	262–323	19.76	0.52	72.46588462	48 (77.4%)	14 (22.6%)
51–102	232–293	18.11	0.5	86.88392308	47 (75.8%)	15 (24.2%)
162–213	232–293	18.52	0.5	60.44236538	47 (75.8%)	15 (24.2%)
162–213	262–323	18.38	0.5	60.44236538	48 (77.4%)	14 (22.6%)
201–252	241–302	18.14	0.5	68.73494231	45 (72.6%)	17 (27.4%)
276–327	232–293	18.45	0.5	72.46588462	47 (75.8%)	15 (24.2%)
312–363	151–212	18.08	0.5	92.69209615	39 (62.9%)	23 (28.1%)
312–363	172–233	18.82	0.5	92.69209615	36 (58.1%)	26 (41.9%)
312–363	211–272	18.67	0.5	92.69209615	48 (77.4%)	14 (22.6%)

[a] Discriminative Power (statistical measure) ranges from 0% (unpredictability) to 100% (predictability). DP values above 50% indicate that the interaction is likely to take place, whereas DPs above 75% represent high-confidence predictions

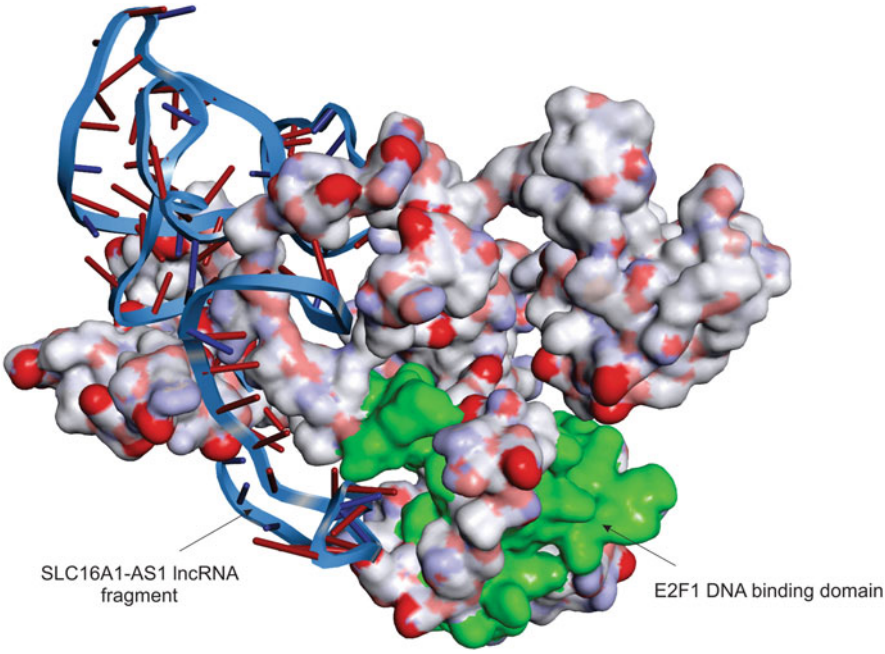


Fig. 11.9 Best docking pose of SLC16A1-AS1 lncRNA fragment (ribbon model) with E2F1 (surface model). DNA binding domain of E2F1 is shown in green color

and lncRNA as identified from the catRAPID fragment tool, the deletions 4–6 were from the loop forming regions that interact with E2F1.

After deletion of the selected segments from the lncRNA, we refolded the lncRNA structure and performed molecular docking with E2F1 with the parameters described above. In all the cases, we keep the lncRNA fragment with a similar length by adding a number of bases, equal to the deleted bases, in both directions. In all the mutagenesis experiments, top interaction poses were analyzed and compared for binding energy between E2F1 and SLC16A1-AS1. For the control case, i.e., lncRNA fragment without deletion, the binding energy of the complex (ΔG) was -35.68 kcal/mol. In case of the 172–175 bases deletion, it formed a more stable complex with E2F1 ($\Delta G = -57.29$ kcal/mol). This is due to the overall increase in the loop regions of the lncRNA after the refolding of the mutated structure. However, the deletion of regions 172–175 and 193–209 together resulted in a weaker complex ($\Delta G = -24.66$ kcal/mol). In case of the deletion of both 172–233 and 262–323 together, no binding regions were identified between E2F1 and the lncRNA by the catRAPID fragmentation tool. In all the cases of in silico mutagenesis, we found both favorable and unfavorable impacts on E2F1 binding affinity, which suggests that this region of SLC16A1-AS1 is important for E2F1 binding. In silico mutagenesis results were experimentally confirmed by the transfection of UMUC-3-KO cells with wild-type SLC16A1-AS1 or the first three deletion mutants, followed

by E2F1 immunoprecipitation, RNA purification, and amplification by qPCR and semiquantitative PCR with primers against SLC16A1-AS1 (Logotheti et al. 2020).

11.6.1.5 Molecular Dynamics Simulation of Best Binding Poses of SLC16A1-AS1 lncRNA and E2F1

The best docked pose was subjected to an MD simulation study using the GROMACS software package (4.5.3) to analyze the interaction stability. The AMBER force field was applied to the complex prior to the MD simulation. The complex was solvated with single point charge (SPC) water molecules (Parrinello and Rahman 1981). The system was further neutralized using proper counter ions by replacing the water molecules to ensure overall charge neutrality of the system. The system was equilibrated by 5000 steps of energy minimization using the steepest descent algorithm, followed by a 100 pico-second (ps) MD equilibrium simulation in constant number of particles, volume, and temperature (NVT) ensemble, with harmonic restraints (20 kcal mol^{-1} $Å^{-2}$) applied to the backbone atoms of the biomolecules. The entire simulation was performed in the isothermal-isobaric ensemble, and both lncRNA fragment and E2F1 were kept unconstrained throughout the simulation run. Temperature and pressure were controlled at 1 atm, and 310 K using a Parrinello-Rahman barostat and V-rescale thermostat respectively as described in (Nosé 1984; Bussi et al. 2007). For the analysis of molecular interactions, a non-bonded cutoff was set to 10 Å and all the electrostatic interactions were calculated using particle mesh Ewald sums (Essmann et al. 1995). Bonds between hydrogen and heavy atoms were constrained at their equilibrium length using the LINCS algorithm (Hess et al. 1997). The production run of 5 ns was performed to study conformational changes during the simulation run time. All trajectories were saved after each 1 ps interval.

We analyzed the root−mean − square deviation (RMSD) of the distance between E2F1- lncRNA fragment and the radius of gyration (Rg) of the complex (Fig. 11.9). To extract the information on the dynamic stability of the docked complex, the root mean square deviation (RMSD) profile of the backbone atoms of the complex was computed with reference to their initial structures for 5 ns. As shown in Fig. 11.10 (left panel), the complex stabilizes at 3 ns. We further analyzed the overall compactness (radius of gyration (Rg) in Fig. 11.10 (middle panel) and the stability of E2F1 in the complex with the lncRNA fragment. Results indicate that after around 3 ns E2F1 attained equilibrium. We also calculated the distance between E2F1 and the lncRNA fragment during the 5 ns production run as shown in Fig. 11.10 (right panel). For this, we calculated the distance between the center of mass (COM) of E2F1 and the lncRNA fragment. Overall, these results indicate that lncRNA SLC16A1-AS1 forms a stable complex with E2F1.

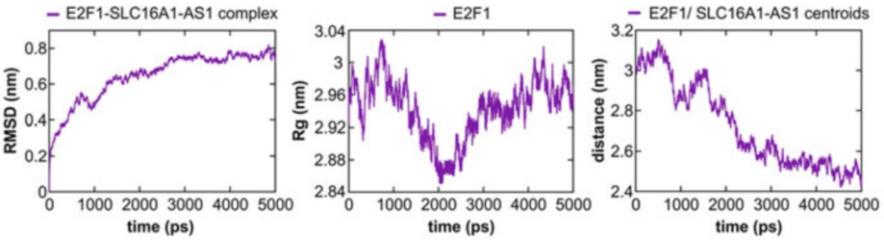


Fig. 11.10 RMSD of docked complex followed a deviation from 0.1 to ~0.7 nm in the initial 2 ns of simulation (left panel) and achieved a significant stability afterwards. Radius of gyration (Rg) (middle panel) values of E2F1 state maintain a relatively steady value of around 2.96 nm. This means that the system shows stability in its folding and achieved its native confirmation. Distance between E2F1 and lncRNA decreases initially and stabilizes at around 3 ns (right panel)

11.6.1.6 Modeling of the SLC16A1-AS1/E2F1 Complex on the Promoter Site of the MCT1 Gene

We further investigated if SLC16A1-AS1 functions in a *cis*-acting manner to regulate transcription of the E2F1 target gene MCT1. For this, we first prepared the 3D structure of the MCT1 promoter region with an extended E2F1 binding motif using the "Build and Edit Nucleic Acid" protocol in the Biovia Discovery Studio. We superimposed the SLC16A1-AS1 fragment in complex with E2F1 to the full-length 3D structure of the lncRNA to prepare a complete receptor molecule and again used the combination of PatchDock and FireDock programs to study the interactions between lncRNA-E2F1 and the MCT1 promoter. Top-ranked poses suggest that the E2F1 DNA binding domain directly interacts with the MCT1 promoter and a part of the lncRNA also interacts with various MCT1 promoter bases on both sites of E2F1 binding domain (Fig. 11.11), suggesting that SLC16A1-AS1 can help stabilizing E2F1 binding to the MCT1 promoter.

We also used the LongTarget tool (He et al. 2015) to identify potential lncRNA-promoter binding sites based on base paring rules and observed a consensus between both sequence and structure level predictions. All the steps presented in this case study are reproducible and can be used to design similar work with other lncRNA–protein interactions.

11.7 Future Directions for the 3D Interactions of ncRNAs

The main issue in the prediction of ncRNA interactions with proteins and other biomolecules is the availability of experimentally determined tertiary structures. Although conventional molecular modeling and molecular docking tools designed for protein–ligand and protein–DNA/RNA interactions are frequently used for ncRNAs as well, the reliability of the results is one of the major concerns specifically for large ncRNA molecules due to their large interaction surface and flexibility of the

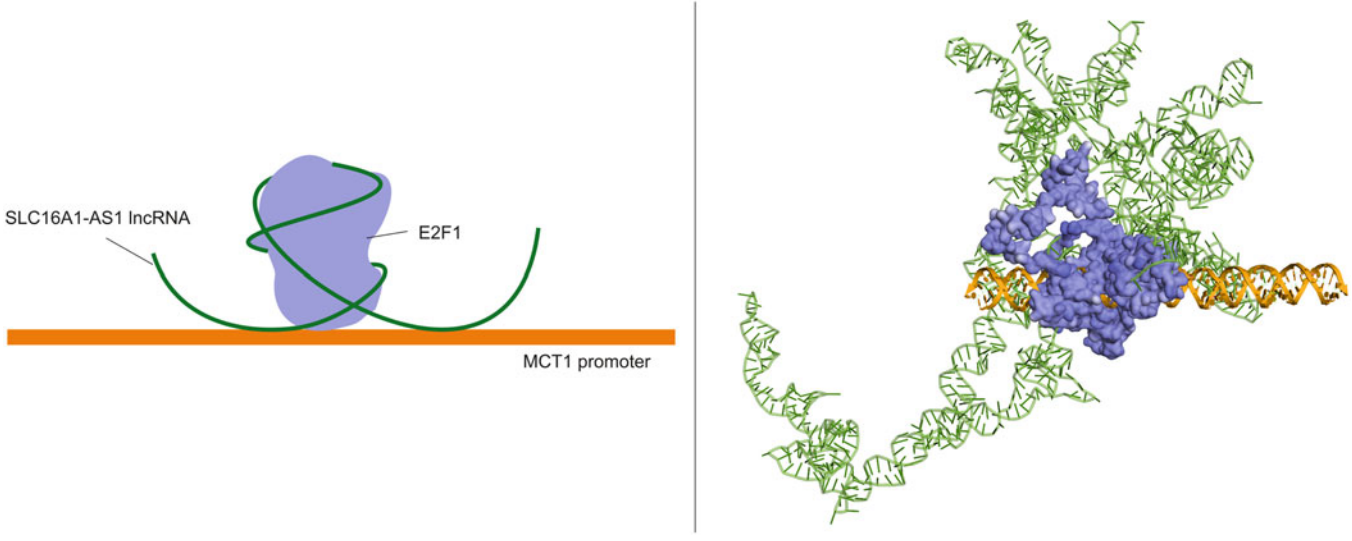


Fig. 11.11 Interaction of the SLC16A1-AS1/E2F1 complex with the MCT1 promoter. Schematic diagram on left suggests that the lncRNA not only binds with the TF but also stabilizes it on the promoter by direct interactions with both sites of the E2F1 binding domain. The 3D structure of the whole complex is shown on the right

molecules. With the research focus shifting toward mechanistic insights of ncRNAs, models of even larger ncRNAs are becoming available, due to the help of experimental data as well as computational prediction pipelines.

Currently available tools for RNA structure modeling achieve good accuracies in predicting topologies and Watson–Crick base pairs for shorter RNA molecules. However, the prediction of large ncRNA structures based on sequence information remains a challenging problem because large ncRNAs form complex tertiary structures that are stabilized by many long-range interactions, non-canonical base pairs, and structural motifs. The prediction of non-canonical interactions (non-Watson–Crick base pairs), which are dominated in case of ncRNA–ncRNA and ncRNA–protein interactions, is also difficult. Advancements in the artificial intelligence-based design of 3D structures and the prediction of interactions at structure level bridge this gap. Further experimental data on non-canonical interactions will help currently available algorithms to predict reliable ncRNA structures and facilitate a mechanistic understanding of their role at the structural level.

References

Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B, Lindah E (2015) GROMACS: high performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX 1–2:19–25

Abulwerdi FA, Xu W, Ageeli AA, Yonkunas MJ, Arun G, Nam H, Schneekloth JS, Dayie TK, Spector D, Baird N, Le Grice SFJ (2019) Selective small-molecule targeting of a triple helix encoded by the Long noncoding RNA, MALAT1. ACS Chem Biol

Adcock SA, McCammon JA (2006) Molecular dynamics: survey of methods for simulating the activity of proteins. Chem Rev 106(5):1589

Agarwal V, Bell GW, Nam JW, Bartel DP (2015) Predicting effective microRNA target sites in mammalian mRNAs. elife 4(August 2015)

Agostini F, Zanzoni A, Klus P, Marchese D, Cirillo D, Tartaglia GG (2013) CatRAPID omics: a web server for large-scale prediction of protein-RNA interactions. Bioinformatics

Akbaripour-Elahabad M, Zahiri J, Rafeh R, Eslami M, Azari M (2016) rpiCOOL: a tool for In Silico RNA-protein interaction detection using random forest. J Theor Biol 402(1):1–8

Alipanahi B, Delong A, Weirauch MT, Frey BJ (2015) Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. Nature Biotechnol 33(8):831–838

Armaos A, Colantoni A, Proietti G, Rupert J, Tartaglia GG (2021) catRAPID omics v2.0: going deeper and wider in the prediction of protein-RNA interactions. Nucleic Acids Res 2(49):72–79

Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, Wang J, Cong Q, Kinch LN, Dustin Schaeffer R, Millán C, Park H, Adams C, Glassman CR, DeGiovanni A, Pereira JH, Rodrigues AV, Van Dijk AA, Ebrecht AC, Opperman DJ, Sagmeister T, Buhlheller C, Pavkov-Keller T, Rathinaswamy MK, Dalwadi U, Yip CK, Burke JE, Christopher Garcia K, Grishin NV, Adams PD, Read RJ, Baker D (2021) Accurate prediction of protein structures and interactions using a three-track neural network. Science 373(6557):871–876

Bandyopadhyay S, Ghosh D, Mitra R, Zhao Z (2015) MBSTAR: multiple instance learning for predicting specific functional binding sites in microRNA targets. Sci Rep 5

Bayly CI, Merz KM, Ferguson DM, Cornell WD, Fox T, Caldwell JW, Kollman PA, Cieplak P, Gould IR, Spellmeyer DC (2002) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 117(19):5179–5197

Bellucci M, Agostini F, Masin M, Tartaglia GG (2011) Predicting protein associations with long noncoding RNAs. Nat Methods

Betel D, Koppal A, Agius P, Sander C, Leslie C (2010) Comprehensive modeling of microRNA targets predicts functional non-conserved and non-canonical sites. Genome Biol 11(8)

Biesiada M, Purzycka KJ, Szachniuk M, Blazewicz J, Adamiak RW (2016) Automated RNA 3D structure prediction with RNA composer. Methods in molecular biology (Clifton NJ) 1490:199–215

Bjørnholt Grønning AG, Doktor TK, Larsen SJ, Spangsberg Petersen US, Holm LL, Bruun GH, Hansen MB, Hartung AM, Baumbach J, Andresen BS (2020) Deep CLIP: predicting the effect of mutations on protein-RNA binding with deep learning. Nucleic Acids Res 48(13):7099–7118

Boniecki MJ, Lach G, Dawson WK, Tomala K, Lukasz P, Soltysinski T, Rother KM, Bujnicki JM (2016) SimRNA: a coarse-grained method for RNA folding simulations and 3D structure prediction. Nucleic Acids Res 44(7)

Bose D, Nahar S, Rai MK, Ray A, Chakraborty K, Maiti S (2015) Selective inhibition of miR-21 by phage display screened peptide. Nucleic Acids Res 43(8):4342–4352

Bozgeyik E, Kocahan S, Temiz E, Bagis H (2021) miR-19a and miR-421 target PCA3 long non-coding RNA and restore PRUNE2 tumor suppressor activity in prostate cancer. Mol Biol Rep:1–13

Bussi G, Donadio D, Parrinello M (2007) Canonical sampling through velocity rescaling. J Chem Phys 126(1):014101

Cech TR, Steitz JA (2014) The noncoding RNA revolution–trashing old rules to forge new ones. Cell 157(1):77–94

Cheatham TE, Cieplak P, Kollman PA (1999) A modified version of the Cornell et al. Force field with improved sugar pucker phases and helical repeat. J Biomol Struct Dyn 16(4):845–862

Chen YC, Sargsyan K, Wright JD, Huang YS, Lim C (2014) Identifying RNA-binding residues based on evolutionary conserved structural and energetic features. Nucleic Acids Res 42(3)

Cheng J, Kapranov P, Drenkow J, Dike S, Brubaker S, Patel S, Long J, Stern D, Tammana H, Helt G, Sementchenko V, Piccolboni A, Bekiranov S, Bailey DK, Ganesh M, Ghosh S, Bell I, Gerhard DS, Gingeras TR (2005) Transcriptional maps of 10 human chromosomes at 5-nucleotide resolution. Science 308(5725):1149–1154

Chipman LB, Pasquinelli AE (2019) miRNA targeting: growing beyond the seed. Trends Genet 35(3):215–222

Christen M, Hünenberger PH, Bakowies D, Baron R, Bürgi R, Geerke DP, Heinz TN, Kastenholz MA, Kräutler V, Oostenbrink C, Peter C, Trzesniak D, van Gunsteren WF (2005) The GROMOS software for biomolecular simulation: GROMOS05. J Comput Chem 26(16): 1719–1751

Chuai G, Ma H, Yan J, Chen M, Hong N, Xue D, Zhou C, Zhu C, Chen K, Duan B, Gu F, Qu S, Huang D, Wei J, Liu Q (2018) Deep CRISPR: optimized CRISPR guide RNA design by deep learning. Genome Biol 19(1)

Cordero P, Das R (2015) Rich RNA structure landscapes revealed by mutate-and-map analysis. PLoS Comput Biol 11(11):e1004473

Cornell WD, Cieplak P, Bayly CI, Gould IR, Merz KM, Ferguson DM, Spellmeyer DC, Fox T, Caldwell JW, Kollman PA (1995) A second generation force field for the simulation of proteins, nucleic acids, and organic molecules. J Am Chem Soc 117(19):5179–5197

Darty K, Denise A, Ponty Y (2009) VARNA: interactive drawing and editing of the RNA secondary structure. Bioinformatics 25(15):1974–1975

Das P, Sercu T, Wadhawan K, Padhi I, Gehrmann S, Cipcigan F, Chenthamarakshan V, Strobelt H, dos Santos C, Chen PY, Yang YY, Tan JPK, Hedrick J, Crain J, Mojsilovic A (2021) Accelerated antimicrobial discovery via deep generative models and molecular dynamics simulations. Nat Biomed Eng 5(6):613–623

Ding F, Sharma S, Chalasani P, Demidov VV, Broude NE, Dokholyan NV (2008) Ab initio RNA folding by discrete molecular dynamics: from structure prediction to folding mechanisms. RNA 14(6):1164–1173

Dominguez C, Boelens R, Bonvin AMJJ (2003) HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. J Am Chem Soc 125(7):1731–1737

Dorn ME, Silva MB, Buriol LS, Lamb LC (2014) Three-dimensional protein structure prediction: methods and computational strategies. Comput Biol Chem 53(PB):251–276

Draper DE, Grilley D, Soto AM (2005) Ions and RNA folding. Annu Rev Biophys Biomol Struct 34:221–243

Duan Y, Wu C, Chowdhury S, Lee MC, Xiong G, Zhang W, Yang R, Cieplak P, Luo R, Lee T, Caldwell J, Wang J, Kollman P (2003) A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J Comput Chem 24(16):1999–2012

Dweep H, Gretz N (2015) MiRWalk2.0: a comprehensive atlas of microRNA-target interactions. Nat Methods 12(8):697

Eichhorn CD, Yang Y, Repeta L, Feigon J (2018) Structural basis for recognition of human 7SK long noncoding RNA by the La-related protein Larp7. Proc Natl Acad Sci U S A 115(28): E6457–E6466

Essmann U, Perera L, Berkowitz ML, Darden T, Lee H, Pedersen LG (1995) A smooth particle mesh Ewald method. J Chem Phys 103(19):8577–8593

Feingold EA, Good PJ, Guyer MS, Kamholz S, Liefer L, Wetterstrand K, Collins FS, Gingeras TR, Kampa D, Sekinger EA, Cheng J, Hirsch H, Ghosh S, Zhu Z, Patel S, Piccolboni A, Yang A, Tammana H, Bekiranov S, Kapranov P, Harrison R, Church G, Struhl K, Ren B, Kim TH, Barrera LO, Qu C, van Calcar S, Luna R, Glass CK, Rosenfeld MG, Guigo R, Antonarakis SE, Birney E, Brent M, Pachter L, Reymond A, Dermitzakis ET, Dewey C, Keefe D, Denoeud F, Lagarde J, Ashurst J, Hubbard T, Wesselink JJ, Castelo R, Eyras E, Myers RM, Sidow A, Batzoglou S, Trinklein ND, Hartman SJ, Aldred SF, Anton E, Schroeder DI, Marticke SS, Nguyen L, Schmutz J, Grimwood J, Dickson M, Cooper GM, Stone EA, Asimenos G, Brudno M, Dutta A, Karnani N, Taylor CM, Kim HK, Robins G, Stamatoyannopoulos G, Stamatoyannopoulos JA, Dorschner M, Sabo P, Hawrylycz M, Humbert R, Wallace J, Yu M, Navas PA, McArthur M, Noble WS, Dunham I, Koch CM, Andrews RM, Clelland GK, Wilcox S, Fowler JC, James KD, Groth P, Dovey OM, Ellis PD, Wraight VL, Mungall AJ, Dhami P, Fiegler H, Langford CF, Carter NP, Vetrie D, Snyder M, Euskirchen G, Urban AE, Nagalakshmi U, Rinn J, Popescu G, Bertone P, Hartman S, Rozowsky J, Emanuelsson O, Royce T, Chung S, Gerstein M, Lian Z, Lian J, Nakayama Y, Weissman S, Stolc V, Tongprasit W, Sethi H, Jones S, Marra M, Shin H, Schein J, Clamp M, Lindblad-Toh K, Chang J, Jaffe DB, Kamal M, Lander ES, Mikkelsen TS, Vinson J, Zody MC, de Jong PJ, Osoegawa K, Nefedov M, Zhu B, Baxevanis AD, Wolfsberg TG, Crawford GE, Whittle J, Holt IE, Vasicek TJ, Zhou D, Luo S, Green ED, Bouffard GG, Margulies EH, Portnoy ME, Hansen NF, Thomas PJ, McDowell JC, Maskeri B, Young AC, Idol JR, Blakesley RW, Schuler G, Miller W, Hardison R, Elnitski L, Shah P, Salzberg SL, Pertea M, Majoros WH, Haussler D, Thomas D, Rosenbloom KR, Clawson H, Siepel A, Kent WJ, Weng Z, Jin S, Halees A, Burden H, Karaoz U, Fu Y, Yu Y, Ding C, Cantor CR, Kingston RE, Dennis J, Green RD, Singer MA, Richmond TA, Norton JE, Farnham PJ, Oberley MJ, Inman DR, McCormick MR, Kim H, Middle CL, Pirrung MC, Fu XD, Kwon YS, Ye Z, Dekker J, Tabuchi TM, Gheldof N, Dostie J, Harvey SC (2004) The ENCODE (ENCyclopedia of DNA elements) project. Science 306(5696):636–640

Feng J, Yang G, Liu Y, Gao Y, Zhao M, Bu Y, Yuan H, Yuan Y, Yun H, Sun M, Gao H, Zhang S, Liu Z, Yin M, Song X, Miao Z, Lin Z, Zhang X (2019) LncRNA PCNAP1 modulates hepatitis B virus replication and enhances tumor growth of liver cancer. Theranostics 9(18):5227–5245

François-Moutal L, Miranda VG, Mollasalehi N, Gokhale V, Khanna M (2021) In silico targeting of the Long noncoding RNA MALAT1. ACS Med Chem Lett 12(6):915–921

Gamazon ER, Im HK, Duan S, Lussier YA, Cox NJ, Dolan ME, Zhang W (2010) ExprTarget: an integrative approach to predicting human microRNA targets. PLoS One 5(10)

Gan HH, Gunsalus KC (2013) Tertiary structure-based analysis of microRNA-target interactions. RNA 19(4):539–551

Gan HH, Gunsalus KC (2015) Assembly and analysis of eukaryotic Argonaute-RNA complexes in microRNA-target recognition. Nucleic Acids Res 43(20):9613–9625

Ghosh S, Talukdar PD, Bhattacharjee A, Giri S, Bhattacharyya NP, Chatterji U (2021) JunD accentuates arecoline-induced disruption of tight junctions and promotes epithelial-to-mesenchymal transition by association with NEAT1 lncRNA. Oncotarget 12(15):1520–1539

Goody D, Gupta SK, Engelmann D, Spitschak A, Marquardt S, Mikkat S, Meier C, Hauser C, Gundlach JP, Egberts JH, Martin H, Schumacher T, Trauzold A, Wolkenhauer O, Logotheti S, Pützer BM (2019) Drug repositioning inferred from E2F1-coregulator interactions studies for the prevention and treatment of metastatic cancers. Theranostics

Gorodkin J, Hofacker IL (2011) From structure prediction to genomic screens for novel non-coding RNAs. PLoS Comput Biol 7(8)

Górska A, Jasiński M, Trylska J (2015) MINT: software to identify motifs and short-range interactions in trajectories of nucleic acids. Nucleic Acids Res 43(17):e114–e114

Gruber AR, Bernhart SH, Lorenz R (2015) The Vienna RNA web services. Methods Mol Biol 1269:307–326

Gruber AR, Lorenz R, Bernhart SH, Neuböck R, Hofacker IL (2008) The Vienna RNA websuite. Nucleic Acids Res

Hassanzadeh HR, Wang MD (2016) Deeper Bind: enhancing prediction of sequence specificities of DNA binding proteins. Proceedings IEEE International Conference on Bioinformatics and Biomedicine 2016:178–183

He J, Tao H, Huang SY (2019) Protein-ensemble-RNA docking by efficient consideration of protein flexibility through homology models. Bioinformatics (Oxford, England) 35(23): 4994–5002

He S, Zhang H, Liu H, Zhu H (2015) Long target: a tool to predict lncRNA DNA-binding motifs and binding sites via Hoogsteen base-pairing analysis. Bioinformatics (Oxford, England) 31(2): 178–186

Helwak A, Kudla G, Dudnakova T, Tollervey D (2013) Mapping the human miRNA interactome by CLASH reveals frequent noncanonical binding. Cell 153(3):654–665

Hess B, Bekker H, Berendsen HJC, Fraaije JGEM (1997) LINCS: a linear constraint solver for molecular simulations. J Comput Chem 18(12):1463–1472

Hiller M, Pudimat R, Busch A, Backofen R (2006) Using RNA secondary structures to guide sequence motif finding towards single-stranded regions. Nucleic Acids Res 34(17)

Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, De Groot BL, Grubmüller H, MacKerell AD (2017) CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat Methods 14(1):71–73

Iwakiri J, Hamada M, Asai K, Kameda T (2016) Improved accuracy in RNA-protein rigid body docking by incorporating force field for molecular dynamics simulation into the scoring function. J Chem Theory Comput 12(9):4688–4697

Jha A, Saha S, Ayasolla K, Vashistha H, Malhotra A, Skorecki K, Singhal PC (2020) MiR193a modulation and podocyte phenotype. Cell 9(4):1004

Jiang H, Sheong FK, Zhu L, Gao X, Bernauer J, Huang X (2015) Markov state models reveal a two-step mechanism of miRNA loading into the human Argonaute protein: selective binding followed by structural re-arrangement. PLoS Comput Biol 11(7):e1004404

Jolma A, Zhang J, Mondragón E, Morgunova E, Kivioja T, Laverty KU, Yin Y, Zhu F, Bourenkov G, Morris Q, Hughes TR, Maher LJ, Taipale J (2020) Binding specificities of human RNA-binding proteins toward structured and linear RNA sequences. Genome Res 30(7):962–973

Jorgensen WL, Maxwell DS, Tirado-Rives J (1996) Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. J Am Chem Soc 118(45):11225–11236

Kalvari I, Nawrocki EP, Argasinska J, Quinones-Olvera N, Finn RD, Bateman A, Petrov AI (2018) Non-coding RNA analysis using the Rfam database. Curr Protoc Bioinformatics 62(1):e51

Kalvari I, Nawrocki EP, Ontiveros-Palacios N, Argasinska J, Lamkiewicz K, Marz M, Griffiths-Jones S, Toffano-Nioche C, Gautheret D, Weinberg Z, Rivas E, Eddy SR, Finn RD, Bateman A, Petrov AI (2021) Rfam 14: expanded coverage of metagenomic, viral and microRNA families. Nucleic Acids Res 49(D1):D192–D200

Kazan H, Ray D, Chan ET, Hughes TR, Morris Q (2010) RNAcontext: a new method for learning the sequence and structure binding preferences of RNA-binding proteins. PLoS Comput Biol 6(7):28

Kerpedjiev P, Hammer S, Hofacker IL (2015) Forna (force-directed RNA): simple and effective online RNA secondary structure diagrams. Bioinformatics (Oxford, England) 31(20): 3377–3379

Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E (2007) The role of site accessibility in microRNA target recognition. Nat Genet 39(10):1278–1284

Klausen MS, Jespersen MC, Nielsen H, Jensen KK, Jurtz VI, Sønderby CK, Sommer MOA, Winther O, Nielsen M, Petersen B, Marcatili P (2019) NetSurfP-2.0: improved prediction of protein structural features by integrated deep learning. Proteins: Structure, Function and Bioinformatics 87(6):520–527

Koralewska N, Szczepanska A, Ciechanowska K, Wojnicka M, Pokornowska M, Milewski MC, Gudanis D, Baranowski D, Nithin C, Bujnicki JM, Gdaniec Z, Figlerowicz M, Kurzynska-Kokorniak A (2021) RNA and DNA G-quadruplexes bind to human dicer and inhibit its activity. Cell Mol Life Sci 78(7):3709–3724

Kubota M, Tran C, Spitale RC (2015) Progress and challenges for chemical probing of RNA structure inside living cells. Nat Chem Biol 11(12):933–941

Kuhlman B, Bradley P (2019) Advances in protein structure prediction and design. Nat Rev Mol Cell Biol 20(11):681–697

Lam JH, Li Y, Zhu L, Umarov R, Jiang H, Héliou A, Sheong FK, Liu T, Long Y, Li Y, Fang L, Altman RB, Chen W, Huang X, Gao X (2019) A deep learning framework to predict binding preference of RNA constituents on protein surface. Nat Commun 10(1)

Li S, Dong F, Wu Y, Zhang S, Zhang C, Liu X, Jiang T, Zeng J (2017) A deep boosting based approach for capturing the sequence binding preferences of RNA-binding proteins from high-throughput CLIP-seq data. Nucleic Acids Res 45(14):e129

Li Y, Ren Y, Wang Y, Tan Y, Wang Q, Cai J, Zhou J, Yang C, Zhao K, Yi K, Jin W, Wang L, Liu M, Yang J, Li M, Kang C (2019) A compound AC1q3qWB selectively disrupts HOTAIR-mediated recruitment of PRC2 and enhances cancer therapy of DzNEP. Theranostics 9(16): 4608–4623

Liu H, Yue D, Chen Y, Gao SJ, Huang Y (2010) Improving performance of mammalian microRNA target prediction. BMC Bioinformatics 11

Liu Y, Gong W, Zhao Y, Deng X, Zhang S, Li C (2021) aPRBind: protein-RNA interface prediction by combining sequence and I-TASSER model-based structural features learned with convolutional neural networks. Bioinformatics (Oxford, England) 37(7):937–942

Logotheti S, Marquardt S, Gupta SK, Richter C, Edelhäuser BAH, Engelmann D, Brenmoehl J, Söhnchen C, Murr N, Alpers M, Singh KP, Wolkenhauer O, Heckl D, Spitschak A, Pützer BM (2020) LncRNA-SLC16A1-AS1 induces metabolic reprogramming during bladder cancer progression AS target and co-activator of E2F1. Theranostics 10(21):9620–9643

Lorenz R, Bernhart SH, Höner ZU, Siederdissen C, Tafer H, Flamm C, Stadler PF, Hofacker IL (2011) ViennaRNA Package 2.0. Algorithms Mol Biol 6:26

Lorenz R, Wolfinger MT, Tanzer A, Hofacker IL (2016) Predicting RNA secondary structures from sequence and probing data. Methods 103:86–98

Loughrey D, Watters KE, Settle AH, Lucks JB (2014) SHAPE-Seq 2.0: systematic optimization and extension of high-throughput chemical probing of RNA secondary structure with next generation sequencing. Nucleic Acids Res 42(21)

Lu Q, Ren S, Lu M, Zhang Y, Zhu D, Zhang X, Li T (2013) Computational prediction of associations between long non-coding RNAs and proteins. BMC Genomics 14(1):651

Lu C, Wu C, Ghoreishi D, Chen W, Wang L, Damm W, Ross GA, Dahlgren MK, Russell E, Von Bargen CD, Abel R, Friesner RA, Harder ED (2021) OPLS4: improving force field accuracy on challenging regimes of chemical space. J Chem Theory Comput 17(7):4291–4300

Lu X, Ding Y, Bai Y, Li J, Zhang G, Wang S, Gao W, Xu L, Wang H (2020) Detection of allosteric effects of lncRNA secondary structures altered by SNPs in human diseases. Frontiers in Cell and Developmental Biology 8:242

Macindoe G, Mavridis L, Venkatraman V, Devignes MD, Ritchie DW (2010) HexServer: an FFT-based protein docking server powered by graphics processors. Nucleic Acids Res 38. (Web Server issue)

MacKerell AD, Bashford D, Bellott M, Dunbrack RL, Evanseck JD, Field MJ, Fischer S, Gao J, Guo H, Ha S, Joseph-McCarthy D, Kuchnir L, Kuczera K, Lau FT, Mattos C, Michnick S, Ngo T, Nguyen DT, Prodhom B, Reiher WE, Roux B, Schlenkrich M, Smith JC, Stote R, Straub J, Watanabe M, Wiórkiewicz-Kuczera J, Yin D, Karplus M (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. J Phys Chem B 102(18): 3586–3616

Mahernia S, Hassanzadeh M, Sarvari S, Amanlou M (2020) Targeting the microRNA binding domain of argonaute 2: rational inhibitor design and study of mutation effects on protein-ligand interaction. J Biomol Struct Dyn 38(16):4710–4717

Mallick B, Sharma AR, Lee SS, Chakraborty C (2019) Understanding the molecular interaction of human argonaute-2 and miR-20a complex: a molecular dynamics approach. J Cell Biochem 120(12):19915–19924

Mashiach E, Schneidman-Duhovny D, Andrusier N, Nussinov R, Wolfson HJ (2008) Fire dock: a web server for fast interaction refinement in molecular docking. Nucleic Acids Res 36(Web Server issue):W229–W232

Mathews DH (2004) Using an RNA secondary structure partition function to determine confidence in base pairs predicted by free energy minimization. RNA 10(8):1178–1190

Mathews DH, Moss WN, Turner DH (2010) Folding and finding RNA secondary structure. Cold Spring Harb Perspect Biol 2(12):a003665

Maticzka D, Lange SJ, Costa F, Backofen R (2014) Graph Prot: modeling binding preferences of RNA-binding proteins. Genome Biol 15(1)

Meier C, Spitschak A, Abshagen K, Gupta S, Mor JM, Wolkenhauer O, Haier J, Vollmar B, Alla V, Pützer BM (2014) Association of RHAMM with E2F1 promotes tumour cell extravasation by transcriptional up-regulation of fibronectin. J Pathol 234(3):351–364

Miao Y, Feher VA, McCammon JA (2015) Gaussian accelerated molecular dynamics: unconstrained enhanced sampling and free energy calculation. J Chem Theory Comput 11(8): 3584–3595

Morris GM, Ruth H, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ (2009) Software news and updates auto Dock4 and AutoDockTools4: automated docking with selective receptor flexibility. J Comput Chem 30(16):2785–2791

Muppirala UK, Honavar VG, Dobbs D (2011) Predicting RNA-protein interactions using only sequence information. BMC bioinformatics 12(1)

Naderi N, Namvar A, Amani N, Nasoohi N, Bolhassani A (2018) Analysis of long non-coding RNA expression in hemophilia A patients. Hematology 24(1):255–262. https://doi.org/10.1080/16078454.2018.1560934

Nosé S (1984) A unified formulation of the constant temperature molecular dynamics methods. J Chem Phys 81(1):511–519

Orenstein Y, Wang Y, Berger B (2016) RCK: accurate and efficient inference of sequence- and structure-based protein-RNA binding models from RNAcompete data. Bioinformatics (Oxford, England) 32(12):i351–i359

Palermo G, Casalino L, Magistrato A, Andrew McCammon J (2019) Understanding the mechanistic basis of non-coding RNA through molecular dynamics simulations. J Struct Biol 206(3):267

Palermo G, Chen JS, Ricci CG, Rivalta I, Jinek M, Batista VS, Doudna JA, McCammon JA (2018) Key role of the REC lobe during CRISPR-Cas9 activation by 'sensing', 'regulating', and 'locking' the catalytic HNH domain. Q Rev Biophys 51:e91

Pan X, Rijnbeek P, Yan J, Shen HB (2018) Prediction of RNA-protein sequence and structure binding preferences using deep convolutional and recurrent neural networks. BMC Genomics 19(1)

Pan X, Shen HB (2018) Predicting RNA-protein binding sites and motifs through combining local and global deep convolutional neural networks. Bioinformatics (Oxford, England) 34(20): 3427–3436

Pappu RV, Hart RK, Ponder JW (1998) Analysis and application of potential energy smoothing and search methods for global optimization. J Phys Chem B 102(48):9725–9742

Parisien M, Major F (2008) The MC-fold and MC-Sym pipeline infers RNA structure from sequence data. Nature 452(7183):51–55

Parker JS, Parizotto EA, Wang M, Roe SM, Barford D (2009) Enhancement of the seed-target recognition step in RNA silencing by a PIWI/MID domain protein. Mol Cell 33(2):204–214

Parrinello M, Rahman A (1981) Polymorphic transitions in single crystals: a new molecular dynamics method. J Appl Phys 52(12):7182–7190

Petrov AI, Kay SJE, Gibson R, Kulesha E, Staines D, Bruford EA, Wright MW, Burge S, Finn RD, Kersey PJ, Cochrane G, Bateman A, Griffiths-Jones S, Harrow J, Chan PP, Lowe TM, Zwieb CW, Wower J, Williams KP, Hudson CM, Gutell R, Clark MB, Dinger M, Quek XC, Bujnicki JM, Chua NH, Liu J, Wang H, Skogerbø G, Zhao Y, Chen R, Zhu W, Cole JR, Chai B, Da Huang H, Huang HY, Cherry JM, Hatzigeorgiou A, Pruitt KD (2015) RNAcentral: an international database of ncRNA sequences. Nucleic Acids Res 43(Database issue):D123–D129

Philip M, Chen T, Tyagi S (2021) A survey of current resources to study lncrna-protein interactions. Non-coding RNA 7(2):33

Pierce BG, Wiehe K, Hwang H, Kim BH, Vreven T, Weng Z (2014) ZDOCK server: interactive docking prediction of protein-protein complexes and symmetric multimers. Bioinformatics 30(12):1771–1773

Quillet A, Saad C, Ferry G, Anouar Y, Vergne N, Lecroq T, Dubessy C (2020) Improving bioinformatics prediction of microRNA targets by ranks aggregation. Front Genet 10:1330

Rivas E, Eddy SR (2000) Secondary structure alone is generally not statistically significant for the detection of noncoding RNAs. Bioinformatics 16(7):583–605

Rouskin S, Zubradt M, Washietl S, Kellis M, Weissman JS (2013) Genome-wide probing of RNA structure reveals active unfolding of mRNA structures in vivo. Nature 505(7485):701–705

Salerno D, Chiodo L, Alfano V, Floriot O, Cottone G, Paturel A, Pallocca M, Plissonnier ML, Jeddari S, Belloni L, Zeisel M, Levrero M, Guerrieri F (2020) Hepatitis B protein HBx binds the DLEU2 lncRNA to sustain cccDNA and host cancer-related gene transcription. Gut 69(11): 2016–2024

Sarah D, Davis CA, Angelika M, Alex D, Timo L, Ali M, Andrea T, Julien L, Wei L, Felix S (2012) Landscape of transcription in human cells. Nature 489(7414):101–108

Schirle NT, Sheu-Gruttadauria J, MacRae IJ (2014) Structural basis for microRNA targeting. Science 346(6209):608–613

Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ (2005) PatchDock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res 33:363–367

Senior AW, Evans R, Jumper J, Kirkpatrick J, Sifre L, Green T, Qin C, Žídek A, Nelson AWR, Bridgland A, Penedones H, Petersen S, Simonyan K, Crossan S, Kohli P, Jones DT, Silver D, Kavukcuoglu K, Hassabis D (2020) Improved protein structure prediction using potentials from deep learning. Nature 577(7792):706–710

Sheu-Gruttadauria J, Xiao Y, Gebert LF, MacRae IJ (2019) Beyond the seed: structural basis for supplementary micro RNA targeting by human Argonaute2. EMBO J 38(13):e101153

Shindyalov IN, Bourne PE (2001) A database and tools for 3-D protein structure comparison and alignment using the combinatorial extension (CE) algorithm. Nucleic Acids Res 29(1):228–229

Sponer J, Bussi G, Krepl M, Banas P, Bottaro S, Cunha RA, Gil-Ley A, Pinamonti G, Poblete S, Jurečka P, Walter NG, Otyepka M (2018) RNA structural dynamics as captured by molecular simulations: a comprehensive overview [online]. Chemical Rev 118(8):4177–4338. https://doi.org/10.1021/acs.chemrev.7b00427. [Accessed 19 Feb 2022]

Šponer J, Krepl M, Banáš P, Kührová P, Zgarbová M, Jurečka P, Havrila M, Otyepka M (2017) How to understand atomistic molecular dynamics simulations of RNA and protein-RNA complexes? Wiley interdisciplinary reviews RNA 8(3)

Sun M, Wang X, Zou C, He Z, Liu W, Li H (2016) Accurate prediction of RNA-binding protein residues with two discriminative structural descriptors. BMC bioinformatics 17(1)

Suresh V, Liu L, Adjeroh D, Zhou X (2015) RPI-Pred: predicting ncRNA-protein interaction using sequence and structural information. Nucleic Acids Res 43(3):1370–1379

Suyal S, Singh MP, Shekhar H, Srivastava S (2018) In silico screening of proteins targeting circulating miRNAs for improved diagnosis of multiple myeloma. Biochem Biophys Res Commun 497(2):577–582

Talkish J, May G, Lin Y, Woolford JL, McManus CJ (2014) Mod-seq: high-throughput sequencing for chemical probing of RNA structure. RNA 20(5):713

Townshend, R.J.L., Eismann, S., Watkins, A.M., Rangan, R., Karelina, M., Das, R., and Dror, R. O., 2021. Geometric deep learning of RNA structure. Science (New York, N.Y.), 373 (6558), 1047–1051

Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, Darian E, Guvench O, Lopes P, Vorobyov I, Mackerell AD Jr (2010) CHARMM general force field: a force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. J Comput Chem 31(4):671–690

Vejnar CE, Zdobnov EM (2012) MiRmap: comprehensive prediction of microRNA target repression strength. Nucleic Acids Res 40(22):11673–11683

Wang J, Cieplak P, Kollman PA (2000) How well does a restrained electrostatic potential (RESP) model perform in calculating conformational energies of organic and biological molecules? Keywords: additive force field; nonadditive force field; restrained electrostatic potential (RESP); torsional angle parameterization. J Comput Chem 21(12):1049–1074

Wang Y, Alla V, Goody D, Gupta SK, Spitschak A, Wolkenhauer O, Pützer BM, Engelmann D (2016) Epigenetic factor EPC1 is a master regulator of DNA damage response by interacting with E2F1 to silence death and activate metastasis-related gene signatures. Nucleic Acids Res 44(1)

Wang Y, Li Y, Ma Z, Yang W, Ai C (2010) Mechanism of microRNA-target interaction: molecular dynamics simulations and thermodynamics analysis. PLoS Comput Biol 6(7):5

Wang Y, Sheng G, Juranek S, Tuschl T, Patel DJ (2008) Structure of the guide-strand-containing argonaute silencing complex. Nature 456(7219):209–213

Wang J, Wang J, Huang Y, Xiao Y (2019) 3dRNA v2.0: an updated web server for RNA 3D structure prediction. Int J Mol Sci 20(17):4116

Washietl S, Pedersen JS, Korbel JO, Stocsits C, Gruber AR, Hackermüller J, Hertel J, Lindemeyer M, Reiche K, Tanzer A, Ucla C, Wyss C, Antonarakis SE, Denoeud F, Lagarde J, Drenkow J, Kapranov P, Gingeras TR, Guigó R, Snyder M, Gerstein MB, Reymond A, Hofacker IL, Stadler PF (2007) Structured RNAs in the ENCODE selected regions of the human genome. Genome Res 17(6):852–864

Weeks KM (2010) Advances in RNA structure analysis by chemical probing. Curr Opin Struct Biol 20(3):295–304

Workman C, Krogh A (1999) No evidence that mRNAs have lower folding free energies than random sequences with the same dinucleotide distribution. Nucleic Acids Res 27(24):4816

Wu Q, Peng Z, Zhang Y, Yang J (2018) COACH-D: improved protein-ligand binding sites prediction with refined ligand-binding poses through molecular docking. Nucleic Acids Res 46(W1):W438–W442

Xia Y, Xia CQ, Pan X, Shen HB (2021) GraphBind: protein structural context embedded rules learned by hierarchical graph neural networks for recognizing nucleic-acid-binding residues. Nucleic Acids Res 49(9):E51

Xie J, Zheng J, Hong X, Tong X, Liu S (2020) PRIME-3D2D is a 3D2D model to predict binding sites of protein-RNA interaction. Communications biology 3(1)

Xu Y, Vanommeslaeghe K, Aleksandrov A, MacKerell AD, Nilsson L (2016) Additive CHARMM force field for naturally occurring modified ribonucleotides. J Comput Chem 37(10):896

Yan J, Friedrich S, Kurgan L (2016) A comprehensive comparative review of sequence-based predictors of DNA- and RNA-binding residues. Brief Bioinform 17(1):88–105

Yan J, Kurgan L (2017) DRNApred, fast sequence-based method that accurately predicts and discriminates DNA- and RNA-binding residues. Nucleic Acids Res 45(10)

Yan Y, Zhang D, Zhou P, Li B, Huang SY (2017) HDOCK: a web server for protein-protein and protein-DNA/RNA docking based on a hybrid strategy. Nucleic Acids Res 45(W1):W365–W373

Yang J, Roy A, Zhang Y (2013) Protein-ligand binding site recognition using complementary binding-specific substructure comparison and sequence profile alignment. Bioinformatics (Oxford, England) 29(20):2588–2595

Yeh CC, Luo JL, Nhut Phan N, Cheng YC, Chow LP, Tsai MH, Chuang EY, Lai LC (2018) Different effects of long noncoding RNA NDRG1-OT1 fragments on NDRG1 transcription in breast cancer cells under hypoxia. RNA Biol 15(12):1487–1498. https://doi.org/10.1080/15476286.2018.1553480

Zadeh JN, Steenberg CD, Bois JS, Wolfe BR, Pierce MB, Khan AR, Dirks RM, Pierce NA (2011) NUPACK: analysis and design of nucleic acid systems. J Comput Chem 32(1):170–173

Zhang X, Liu S, Tramontano A (2017) RBPPred: predicting RNA-binding proteins from sequence using SVM. Bioinformatics (Oxford, England) 33(6):854–862

Zhang Y, Wang J, Xiao Y (2020) 3dRNA: building RNA 3D structure with improved template library. Comput Struct Biotechnol J 18:2416–2423

Zheng J, Hong X, Xie J, Tong X, Liu S (2020) P3DOCK: a protein-RNA docking webserver based on template-based and template-free docking. Bioinformatics (Oxford, England) 36(1):96–103

Zheng J, Kundrotas PJ, Vakser IA, Liu S (2016) Template-based modeling of protein-RNA interactions. PLoS Comput Biol 12(9)

Zuker M (2003) Mfold web server for nucleic acid folding and hybridization prediction. Nucleic Acids Res 31(13):3406–3415