Cellular and Molecular Life Sciences

REVIEW

Principles of miRNA–mRNA interactions: beyond sequence complementarity

Fabian Afonso-Grunz[1,2] · **Sören Müller**[1,2]

Abstract MicroRNAs (miRNAs) are small non-coding RNAs that post-transcriptionally regulate gene expression by altering the translation efficiency and/or stability of targeted mRNAs. In vertebrates, more than 50 % of all protein-coding RNAs are assumed to be subject to miRNA-mediated control, but current high-throughput methods that reliably measure miRNA–mRNA interactions either require prior knowledge of target mRNAs or elaborate preparation procedures. Consequently, experimentally validated interactions are relatively rare. Furthermore, in silico prediction based on sequence complementarity of miRNAs and their corresponding target sites suffers from extremely high false positive rates. Apparently, sequence complementarity alone is often insufficient to reflect the complex post-transcriptional regulation of mRNAs by miRNAs, which is especially true for animals. Therefore, combined analysis of small non-coding and protein-coding RNAs is indispensable to better understand and predict the complex dynamics of miRNA-regulated gene expression. Single-nucleotide polymorphisms (SNPs) and alternative polyadenylation (APA) can affect miRNA binding of a given transcript from different individuals and tissues, and especially APA is currently emerging as a major factor that contributes to variations in miRNA–mRNA interplay in animals. In this review, we focus on the influence of APA and SNPs on miRNA-mediated gene regulation and discuss the computational approaches that take these mechanisms into account.

F. Afonso-Grunz and S. Müller contributed equally to the publication.

✉ Fabian Afonso-Grunz
FGrunz@rz.uni-frankfurt.de

1 GenXPro GmbH, Frankfurt Innovation Center Biotechnology, Altenhöferallee 3, 60438 Frankfurt am Main, Germany

2 Institute for Molecular BioSciences, Goethe University Frankfurt am Main, Max-von-Laue-Str. 9, 60438 Frankfurt am Main, Germany

Introduction

Among the diverse classes of functional small non-coding RNA (small ncRNA), microRNAs (miRNAs) represent important regulators of gene expression that are expressed in plants, fungi, animals, and unicellular eukaryotes as well as viruses [1–6]. The typically 18–26 nucleotide long miRNAs primarily target messenger RNAs (mRNAs), but binding sites are also present in ncRNAs that can act as competing endogenous RNAs (ceRNAs) [7]. MiRBase represents the most comprehensive database that provides access to genomic miRNA annotations and underlying sequence data along with crosslinks to databases of predicted target genes. The current version (v21) comprises ∼28,600 miRNA entries from 223 different species, including more than 1800 human loci [8–11]. In general, translation of targeted mRNAs is repressed followed by deadenylation and mRNA decay [12–14]. However, in specific cellular conditions such as cell cycle arrest, miRNA binding was also shown to mediate translational up-regulation of target mRNAs [15].

Canonical miRNA biogenesis in plants and animals is a two-step process that initiates with primary miRNA (pri-miRNA) transcription by RNA polymerase II (Pol-II). All pri-miRNAs have a stem-loop structure that is co-

transcriptionally recognized by the Microprocessor complex in animals [16] or by DCL1 (endoribonuclease Dicerlike Dicer-like1) and its paralogs in plants [17]. The Microprocessor is mainly composed of the RNase III Drosha and the double-stranded RNA-binding protein DGCR8 (DiGeorge syndrome critical region 8 gene; also Pasha). DGCR8 recognizes the RNA substrate, while Drosha functions as an endonuclease that generates ∼70 nucleotide long precursor miRNAs (pre-miRNAs) with hairpin structures in a process termed cropping [18, 19]. The pre-miRNA hairpins are subsequently exported to the cytoplasm by Exportin-5 [20], where the second RNase III Dicer mediates cleavage (dicing) of pre-miRNAs into short miRNA duplexes [21, 22]. Only one strand out of the miRNA duplexes (the guide strand) is stably incorporated

into the miRNA-induced silencing complex (miRISC) afterwards (Fig. 1a). Subsequent to loading of the duplex into one of four Argonaute subfamily proteins (AGO1-4) that form the core of each miRISC, the passenger strand (miRNA*) is degraded [23]. Canonical miRNA maturation in plants is also completed by loading of an argonaute complex in the cytoplasm, but the miRNA duplexes are released directly from DCL1 and subsequently exported to the cytoplasm by HASTY [24]. In contrast to mammalian cells, where miRNAs are loaded without much discrimination between different Argonaute subfamily proteins, plant AGO proteins show preferences for specific small RNA classes that are produced by distinct biogenesis pathways [23]. Mirtrons represent a widespread class of intron-derived miRNAs in animals that are generated by a
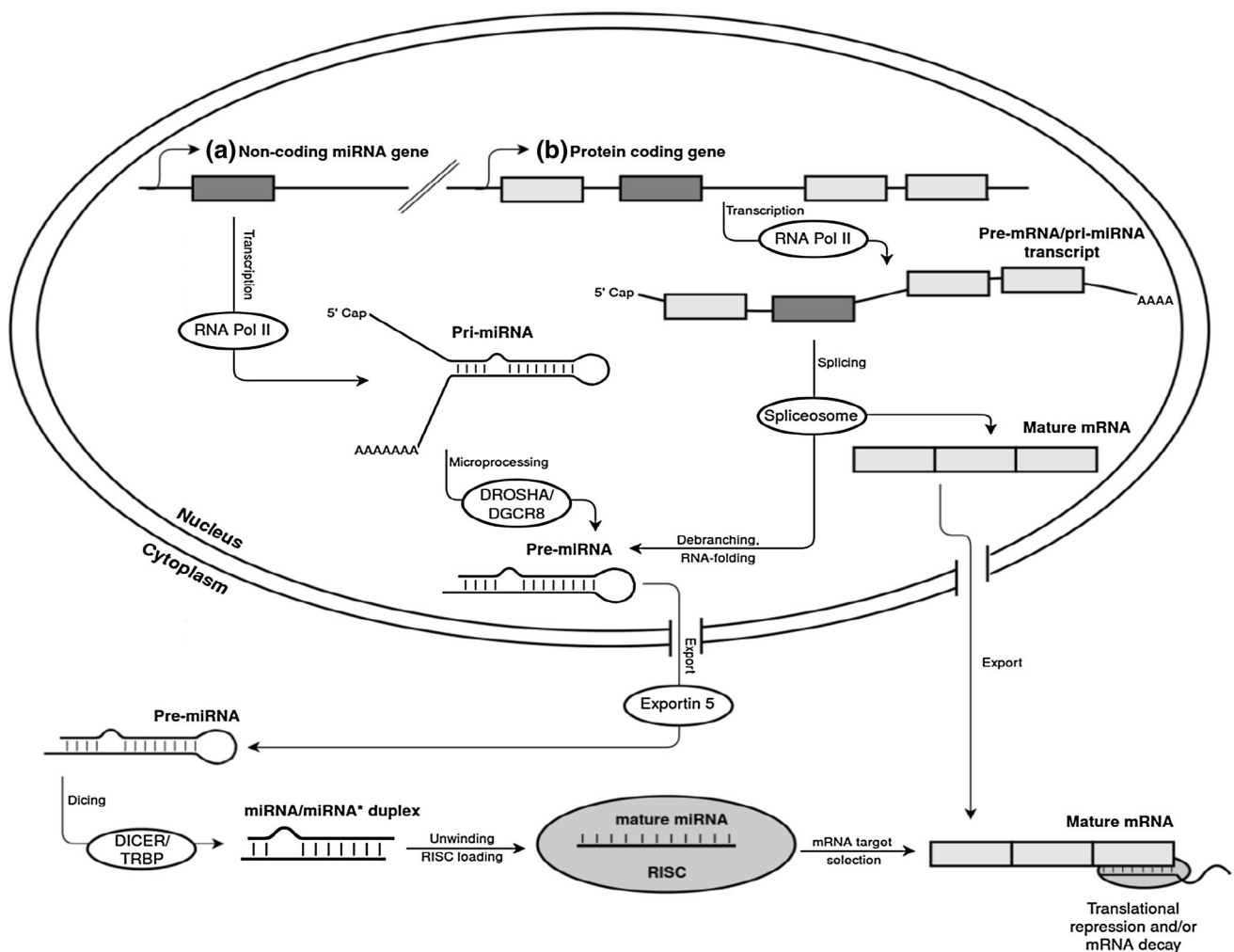

Fig. 1 Canonical miRNA biogenesis and generation of intron-derived miRNAs (mirtrons) in animals. **a** Canonical miRNAs are transcribed independently by Pol-II into pri-miRNAs. Subsequent microprocessing of pri-miRNAs is mediated by the DGCR8/DROSHA complex and produces pre-miRNAs that are exported into the cytoplasm. **b** Mirtrons are transcribed together with a host gene and produced by refolding of debranched intron lariats during mRNA splicing. After export into the cytoplasm, miRNA/miRNA* duplexes are generated by cleavage of the loop via Dicer. The duplexes are loaded into RISC, where only the mature miRNA strand of the duplex is retained, whereas the passenger strand is degraded. In general, partially complementary base pairing of a miRNA seed region with a target mRNA induces translational repression, while perfect base pairing results in siRNA-induced endonucleolytic cleavage

non-canonical biogenesis pathway which bypasses the pri-miRNA processing step by Drosha [25, 26]. Instead of cropping, pre-miRNA-like RNAs are directly produced by refolding of debranched intron lariats during mRNA splicing (Fig. 1b). Since putative mirtrons were also identified by deep sequencing of small ncRNAs in *A. thaliana* and rice, mirtron production is likely to occur also in plants, although to a much lesser extent [27, 28].

Tools for prediction of novel miRNAs from small RNA sequencing data rely on the recognition of a specific read pattern subsequent to mapping to a reference genome [29]. The pattern reflects dicing as the last step of miRNA processing (Fig. 2). Two distinct clusters of accumulated reads cover the miR/miR* loci, separated by a short interval without overlapping reads. This region corresponds to the loop, which is cropped by Dicer and consequently not sequenced. The identified patterns are used for prediction and evaluation of the secondary structure from potential miRNA precursors. The precursors that comply with user-defined criteria are then returned as predicted novel miRNAs.

Partially complementary Watson–Crick base pairing of the miRNA seed region (nucleotides 2–8 from the 5′ end) to a target mRNA typically induces translational repression, while perfect base pairing of the miRNA results in siRNA-induced endonucleolytic cleavage that is mediated by AGO2 in vertebrates [30, 31] or by AGO1 and 10 along with other AGO paralogs in plants [32]. Generally, plant miRNAs show perfect or near-perfect complementarity to their mRNA targets, and consequently, the number of direct targets is lower by at least one order of magnitude compared to animal miRNAs [33]. Canonical plant miRNA target sites are found within the 5′ and 3′ untranslated region (UTR) as well as the coding sequence of mRNAs, suggesting that miRNA-directed regulation in plants is not restricted to a specific RNA context [34]. Against this

backdrop, prediction of miRNA targets in plants requires less elaborate algorithms that mostly rely on sequence complementarity of a given miRNA [35]. Although miRNA binding sites in mammalian cells are also found across the entire mRNA, imperfect 'seed' pairing in 3′ UTRs is considered to mediate post-transcriptional silencing more efficiently than pairing within the 5′ UTR or coding sequence [33, 36, 37]. Consequently, 3′ UTR heterogeneity caused by SNPs or alternative polyadenylation (APA) strongly impacts miRNA-mediated post-transcriptional regulation and has to be taken into account for prediction of miRNA target sites within these cells. In humans, 3′ UTR associated SNPs have been linked to disease susceptibility [38], but the subsets of polymorphisms that have a functional role in regulating gene expression are yet to be defined. Polymorphisms within miRNA binding sites harbor the potential to disrupt miRNA binding or even to introduce novel binding sites in 3′ UTRs, and the biological relevance of these polymorphisms is currently being examined in large case–control studies [39]. APA is a common regulatory mechanism of gene expression that generates mRNAs with distinct 3′ UTRs as well as coding sequences (Fig. 3). More than 70 % of human genes encode primary transcripts that contain multiple polyadenylation sites (PA sites) [40, 41], and a systematic examination of 3′ UTRs produced by APA in murine cells revealed that approximately half of all miRNA target sites are located downstream of the first poly(A) site [42].

In this review we focus on the impact of SNPs and especially APA on miRNA target binding, and furthermore discuss bioinformatical approaches that account for these mechanisms in animals. Quantitative information on APA site usage can be used to extend current prediction algorithms with information on binding site availability to estimate actual interaction efficiencies.

Fig. 2 Mapping pattern of miRNAs. Removal of the pre-miRNA loop by Dicer leads to two separated clusters of reads mapped to the reference genome. The figure is based on http://nbn-resolving.de/urn:nbn:de:bsz:15-qucosa-112876
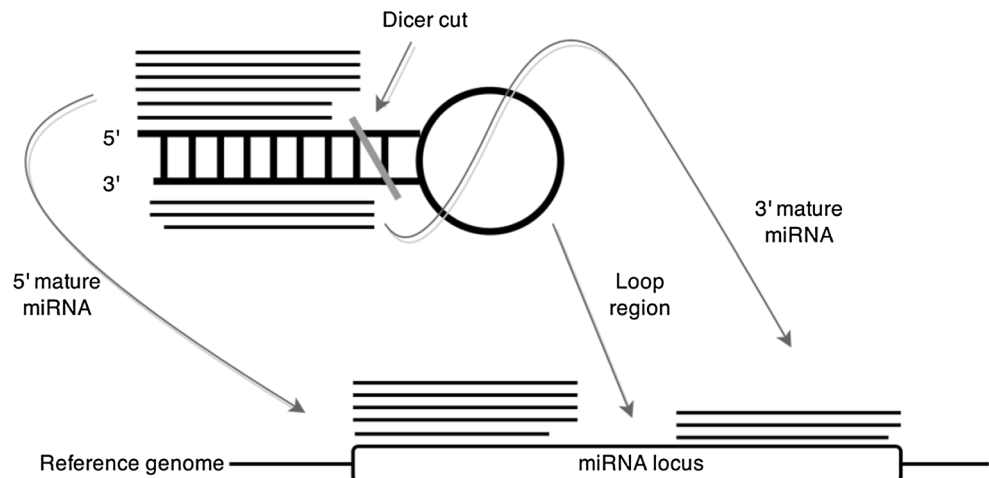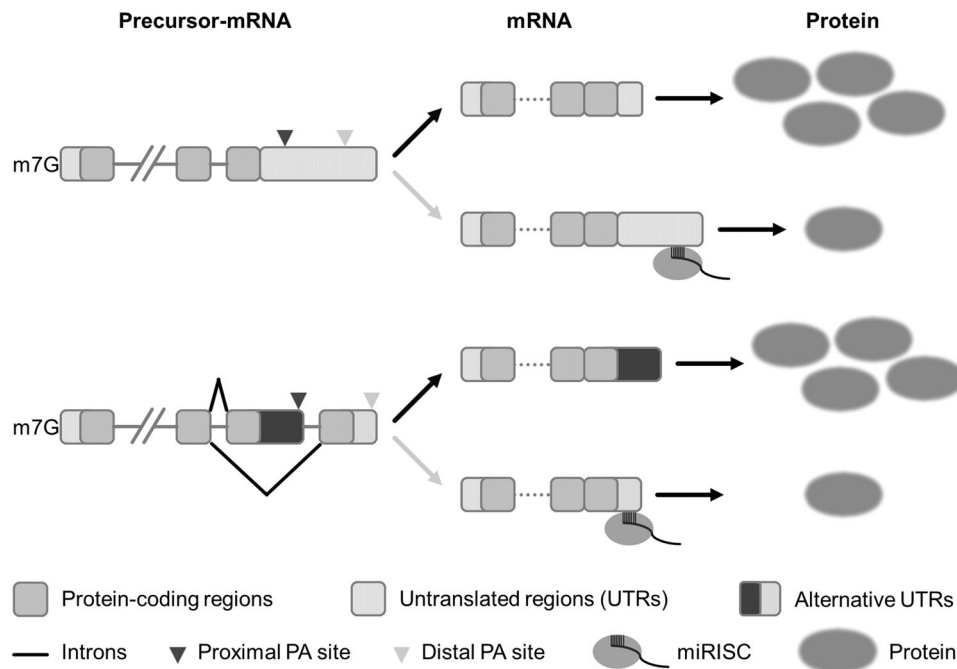
Fig. 3 Alternative polyadenylation (APA) results in formation of distinct mRNA isoforms that are derived from one and the same pre-mRNA. Terminal exons that contain 3′ UTRs with multiple PA sites can give rise to mRNA isoforms that differ in UTR length only (shown on *top*). The length of these tandem UTRs depends on recognition of promotor-proximal PA sites that are linked to shortened 3′ UTR formation or distal PA sites that result in transcription of (nearly) full-length UTRs. Additionally, APA occurs in a splicing-dependent context that comprises the incorporation of alternative terminal exons into the mature mRNA (shown at the *bottom*). In contrast to formation of splicing-independent tandem UTRs, 3′ exon switching affects both the sequence as well as the length of a given 3′ UTR. Shortened or mutually exclusive 3′ UTRs contain different miRNA recognition and protein binding sites that affect the stability, localization, and translation efficiency of an mRNA. Evasion of miRNA-mediated post-transcriptional regulation due to the lack of miRNA binding sites either caused by shortening of tandem UTRs or through incorporation of alternative 3′ terminal exons is typically coupled to increased translation and protein synthesis. Figure adapted from [52]

Alternative polyadenylation and miRNA-mediated post-transcriptional gene regulation

The process of nuclear polyadenylation along with the underlying regulation mechanisms in eukaryotes is summarized in this section. Furthermore, experimental methods for detection of genome-wide APA events are described, and the impact of APA on miRNA binding is discussed based on the data that was generated by these techniques.

Polyadenylation of nascent mRNAs and regulation of PA site usage

Addition of poly(A) tails to the 3′ end of nascent mRNAs in eukaryotic cells occurs as a two-step, co-transcriptional process and depends on the presence of defined poly(A) signals within the pre-mRNA [43, 44]. In mammalian cells, the canonical poly(A) signal AAUAAA is recognized and subsequently bound by CPSF (cleavage and polyadenylation specific factor). Initiation of 3′ end formation is further stimulated by

cooperative binding of CstF (cleavage stimulation factor) to a less defined U/GU-rich downstream element. CFIm, the first of two necessary cleavage factors, additionally assists in poly(A) signal recognition by binding to a third element with the consensus UGUA [45]. Subsequent recruitment of poly(A)-polymerase (PAP) and CFIIm results in endonucleolytic cleavage of the nascent RNA, followed by addition of adenosines to the newly formed 3′ end of the upstream cleavage product via PAP [46]. Together with Symplekin and the C-terminal domain of Pol-II, these proteins constitute the core of the pre-mRNA 3′-end processing machinery [47]. In total, however, ~ 90 proteins contribute to or directly interact with the pre-mRNA 3′-end processing machinery in human cells [48], and some of the corresponding mRNAs are found to be differentially expressed in a context-dependent manner.

A screening of genome-wide APA events in proliferative and arrested human cell lines revealed downregulation of the mRNAs encoding CPSF and CstF, accompanied by global lengthening of 3′ UTRs, during transition from the proliferative to the arrested state [49]. Increased expression of these genes in proliferating cells

was shown to be driven by an enhanced expression of E2F transcription factors, and tightly linked to an increased utilization of proximal PA sites. In contrast, knockdown of CFIm25 resulted in 1450 transcripts with shortened 3′ UTRs in HeLa cells [50]. Decreased expression of CFIm68, another subunit from the CFIm complex, similarly resulted in globally shortened 3′ UTRs [51]. These findings are in line with previous reports of APA events associated with variations in the abundance of core polyadenylation factors (reviewed in [52]), and confirm that altered CFI levels influence PA site usage.

Apart from varying abundances of core members of the 3′-end processing machinery and the presence of cis-acting RNA elements, such as the poly(A) signal, the U/GU-rich downstream element or auxiliary sequences, trans-acting factors like splicing factors and RNA-binding proteins are also involved in the regulation of APA. The U1 small nuclear ribonucleoprotein (U1 snRNP) is usually implicated in mRNA splicing, but was also shown to play a role in APA regulation, where it globally suppresses the usage of distal PA sites [53]. Moderately decreased U1 snRNP levels led to significantly enlarged fractions of mRNAs with proximal PA sites in human, murine and *Drosophila* cells [54]. Di Giammartino and colleagues provide descriptions for several other splicing factors that influence APA including Nova2, PTB (polypyrimidine tract binding protein), hnRNP H (heterogeneous nuclear Ribonucleoprotein H), SRm160 and U2AF65. Besides trans-acting factors, increasing evidence suggests that APA is regulated on the epigenetic level by DNA methylation, nucleosome positioning and posttranslational histone modifications [52]. A systematic analysis of human nucleosome occupancy patterns revealed that highly used proximal PA sites exhibit higher upstream nucleosome levels and Pol-II accumulation than lowly used sites, indicating that the nucleosomes positioned upstream of proximal PA sites influence recognition of these sites by decelerated transcription [55].

Experimental methods for genome-wide identification of PA sites

Techniques that allow for global detection of novel APA events include 3′ end sequencing approaches like Poly(A) site sequencing (PAS-Seq [56]), PolyA-seq [57], poly(A)-position profiling by sequencing (3P-Seq [58]), massive analysis of cDNA ends (MACE [59]), as well as the more recently published 3′ region extraction and deep sequencing (3′READS) protocol [60]. In comparison to RNA-Seq, these methods generate exactly one read out of the 3′ end of each mRNA, which allows for accurate 3′ UTR isoform quantification. PolyA-seq, PAS-seq, and MACE produce oligo(dT)-primed reads proximal to

(0–400 nucleotides) or even overlapping the poly(A) tail. In contrast, 3P-Seq and 3′READS circumvent conventional oligo(dT) priming during library preparation, because oligo(dT)-based reverse transcription is prone to inadvertent priming of homopolymeric adenosine stretches within the mRNAs. Both, 3P-Seq and 3′READS produce reads that are directly adjacent to the poly(A) tail, and thus allow for filtering of those reads with at least one or two non-genomic 3′-terminal adenine bases (termed poly(A) site supporting reads) for further analysis. In fact, 3P-Seq identified more than 8500 3′ UTR isoforms in *C. elegans* that were missed by standard oligo(dT)-based methods, while 3′READS from different mouse tissues revealed more than 5000 PA sites that were overlooked because of flanking homopolymeric adenosine stretches. Based on 3′READS, the number of murine mRNAs known to exhibit APA increased to almost 80 %.

Although library preparation via MACE relies on oligo(dT)-based reverse transcription, mispriming events during library preparation are minimized by hot priming of the mRNAs [61]. Instead of sequential denaturation and reverse transcription as specified for the PolyA-seq or PAS-seq protocol, hybridization of the oligo(dT) primer and subsequent cDNA synthesis are both performed at elevated temperatures, and without temporary cooling of the samples. The comparison of murine APA events either detected by MACE or by 3P-seq, which is not affected by internal priming, revealed a large overlap between high-confidence PA sites detected by the two methods (83 % for MACE and 85 % for 3P-seq). The proportionally similar percentages of PA sites exclusively detected by one of the two methods indicate an efficient exclusion of false positive PA sites arising from internal priming, and confirm the reliability of MACE alongside 3P-seq and 3′READS for PA site detection.

Impact of APA on miRNA-mediated gene regulation

Transcription of mRNA isoforms with distinct 3′ UTRs modulates the post-transcriptional fate of these mRNAs through inclusion or exclusion of miRNA binding sites (Fig. 3) [62]. Additionally, miRNA-mediated post-transcriptional regulation is affected by the altered accessibility of miRNA binding sites in shorter/longer 3′ UTRs due to secondary structure variations, and by the varying proximity to the translation machinery [36]. In different cell types from zebrafish, approximately ten percent of the predicted miRNA–mRNA interactions are influenced by APA [63]. While the shortest 3′ UTR isoforms are present in the ovaries, a significantly more prevalent use of distal PA sites is found in brain tissues. Globally shortened 3′ UTRs were furthermore identified in early mouse embryonic stem cells [64]. Reprogramming of somatic to induced

pluripotent stem cells is associated with a trend towards shortened 3′ UTRs, whereas reprogramming of spermatogonial cells is linked to 3′ UTR lengthening [65]. Together with the globally lengthened 3′ UTRs in proliferative human cell lines described by Elkon and colleagues [40], these findings reflect that vertebrate cells with an increased proliferative potential generally harbor shorter 3′ UTRs.

Several studies emphasized the implications of APA in human diseases. Deregulated expression of the gene encoding brain-derived neurotrophic factor (BDNF) is associated with several neurodegenerative diseases, such as Huntington's disease [66]. Reduced BDNF levels have been observed in corresponding cell models as well as genetic mouse models, and are likely to contribute to the clinical manifestations of the disease [67]. The UTR of the *BDNF* gene harbors two PA sites: one distal (∼3 KB from UTR start, *BDNF-L*) and one proximal (∼350 bps from UTR start, *BDNF-S*) site [41]. The *BDNF-L* isoform contains ten potential miRNA binding sites, predicted by at least four algorithms, whereas *BDNF-S* only contains six of these. Luciferase reporter assays in human embryonic kidney 293 cells (HEK-293) confirmed the direct interaction of miR-1, miR-10b, miR-155 and miR-191 with *BDNF-L*, while *BDNF-S* interacted only with miR-1 and miR-10b [68]. This is consistent with the observation that the short *BDNF* transcript isoform neither carries a miR-155 nor a miR-191 binding site. Furthermore, after transfection with the miR-1 precursor, luciferase activity was significantly lower for the *BDNF-L* isoform that carries three predicted binding sites for this miRNA, compared to *BDNF-S*, which carries only a single miR-1 binding site [68]. The protein level of BDNF, thus, largely depends on PA site usage associated with altered post-transcriptional regulation of the encoding mRNA by miRNAs.

In 2009, Mayr and Bartel hypothesized that APA might represent a mechanism for genes to escape miRNA-mediated post-transcriptional repression in cancer [69]. Based on Northern blot analysis of the 3′ UTR isoforms from six alternatively polyadenylated target genes in 27 cancer cell lines compared to corresponding normal tissues and non-transformed cell lines, they concluded that 3′ UTR shortening is indeed associated with cancerogenesis. Most cancer cell lines expressed significantly shorter APA isoforms, and across all six genes and all 27 cell lines, 25–70 % of the elevated protein-expression levels could be attributed to a loss of post-transcriptional regulation via miRNAs due to 3′ UTR shortening. In ∼30 % of the cell lines, this regulatory loss could even explain all changes in protein levels. Using the dataset published by Sandberg and colleagues [42], Mayr and Bartel confirmed this trend to be a global mechanism on a genome-wide scale. Moreover, the cancer cell lines showed significantly shorter 3′ UTRs than immortalized cell lines despite a comparable

proliferative potential, indicating that shorter 3′ UTRs are not only associated with increased proliferation but also with malignant transformation [69]. By transfection of NIH3T3 cells with retroviral vectors harboring either the shortest 3′ UTR or the full-length isoform of proto-oncogene *IGF2BP1*, they showed that expression of the short isoform greatly promoted cell transformation. An in-depth analysis of the 3′ UTR isoforms of *IGF2BP1* discloses nine functional PA sites in human HLF cancer cell lines, and reveals a varying number of lacking miRNA binding sites in shortened isoforms (Fig. 4). As shown by Mayr and Bartel, these shortened 3′ UTRs positively affect mRNA stability and result in increased amounts of encoded protein product.

Experimental approaches for detection of miRNA–mRNA interactions

Since the discovery of the first miRNAs in *C. elegans* [70–72], several techniques that allow for identification of miRNA target sites have been established. While miRNA–mRNA interactions in plants are readily detected by identification of degraded target intermediates, experimental methods in animals remain laborious and mostly focused on a predefined set of targets [73]. This section summarizes the current experimental techniques for detection/validation of miRNA–mRNA interactions in plants and animals.

Genome-wide experimental identification of miRNA binding sites in plants

Among the techniques for genome-wide analysis of miRNA-mediated mRNA degradation in plants, the most established approaches are termed parallel analysis of RNA ends (PARE) [74], genome-wide mapping of uncapped and cleaved transcripts (GMUCT) [75], and degradome sequencing [76]. In all three protocols, total RNA isolates are enriched for polyadenylated transcripts followed by 5′ adaptor ligation of an RNA oligonucleotide (GMUCT and PARE) or DNA–RNA hybrid (degradome sequencing). Since capped transcripts are not subject to 5′ adaptor ligation, only uncapped RNA fragments are subsequently reverse-transcribed and amplified. Following sequencing of these RNA fragments, miRNA-mediated cleavage sites are identified by alignment of the 5′ ends of the generated reads to respective reference mRNAs.

The first application of PARE revealed a widespread presence of miRNA target sites across all annotated genes in *A. thaliana* and detected most previously validated targets along with nearly all previously predicted but non-validated targets. Stringent analysis of the potential cleavage sites even allowed for identification of putative
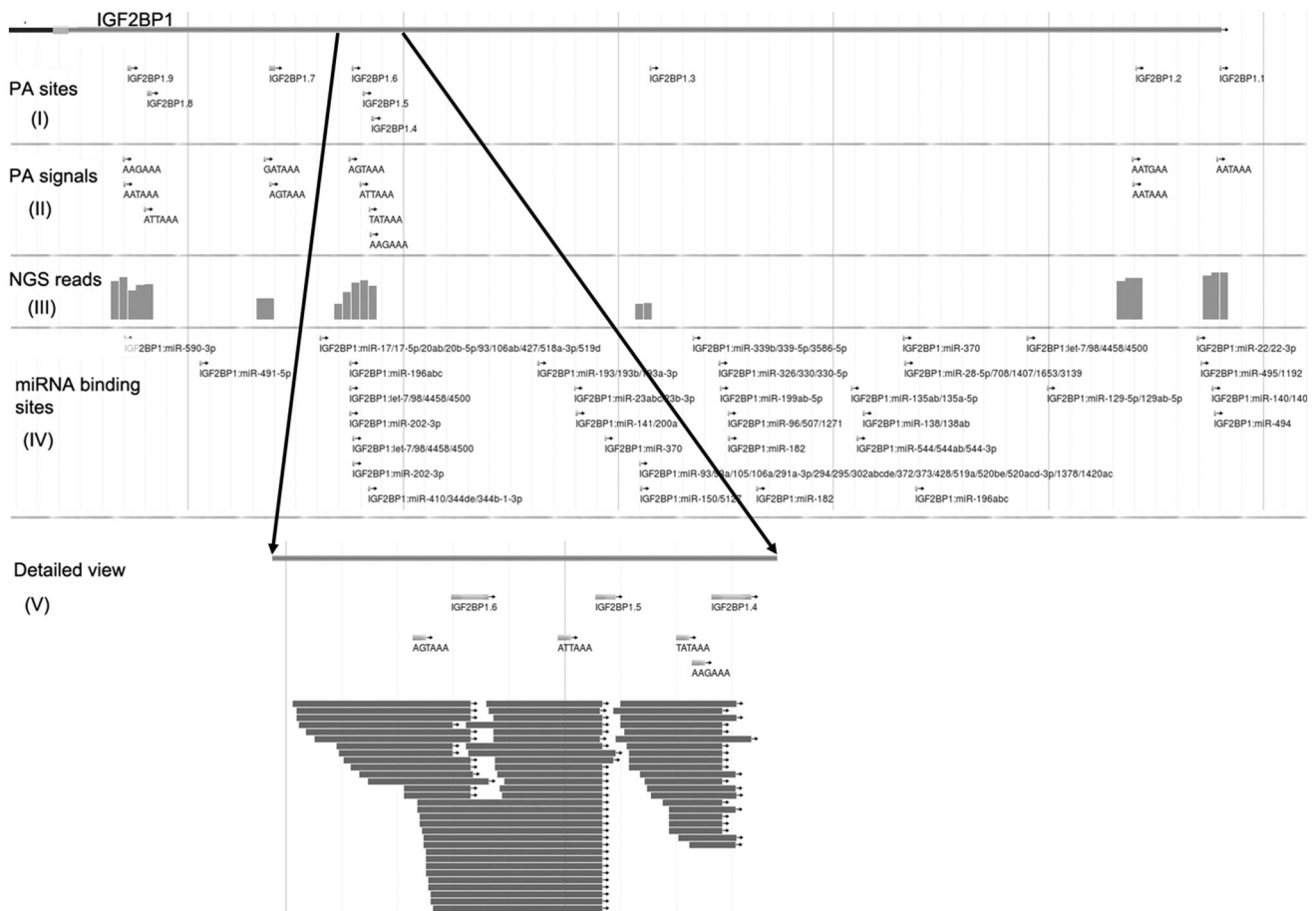
Fig. 4 Visualization of the 3′ UTR from the human *IGF2BP1* transcript by APADB. Nine PA sites (*I*) detected by clustering of end-coordinates of polyA-tail positive NGS reads (*III*) are present in human HLF liver cancer cell lines. Except for one site (*IGF2BP1.3*) each PA site is associated with a closely upstream (∼20 BPs) PA signal (*II*). Only the longest 3′ UTR isoform harbors all predicted TargetScan binding sites (*IV*), while the isoform with the shortest 3′ UTR (*IGF2BP1.9*) does not carry a single predicted miRNA binding site

cleavage sites that were previously unknown, and further validation of some of these targets by 5′ RACE confirmed their presence. At the same time, the first genome-wide screening of miRNA cleavage sites in *A. thaliana* tissues via degradome sequencing resulted in identification of completely novel targets and confirmed previously validated and predicted cleavage sites. To date, genome-wide miRNA–mRNA interaction profiles have been generated for a wide variety of plants including legumes and other important food crops [77, 78]. Additionally, adapted pipelines have been developed that provide tailored computational algorithms for detection of cleaved miRNA targets from degradome data [79, 80]. Only recently, two further developed protocols were published that provide step-by-step instructions for streamlined library preparation for degradome sequencing on Illumina platforms [81, 82]. Together with the adapted pipelines, these techniques provide reliable approaches to globally profile miRNA-mediated cleavage sites also in non-model plants.

Experimental methods for validation of miRNA–mRNA interactions and identification of novel miRNA binding sites in animals

A variety of experimental methods have been developed to detect miRNA–mRNA interactions in animals. The transfection of cell lines with mimic miRNAs or miRNA inhibition by antagomiRs is widely used to measure the effect on mRNA levels within the cells [83, 84]. Even though this technique provides a global profile of changes in mRNA levels via transcriptome profiling by RNA-Seq or similar techniques, direct interactions must be distinguished from indirect interactions and stress responses of the cells have to be discriminated from true responses to miRNA elevation/repression [85]. Reporter assays, on the other hand, allow for identification of direct interactions between a given miRNA and specific mRNA targets [86]. However, these assays are usually limited to a fixed set of interactions, and therefore promising candidate interactions

have to be defined in advance either based on additional experimental data or by in silico prediction of potential miRNA target genes.

Genome-wide methods for detection of miRNA binding sites are based on UV cross-linking in vivo followed by immunoprecipitation of AGO-bound miRNA–mRNA complexes in vitro. HITS-CLIP [87], PAR-CLIP [88], and CLASH [89] represent the most recent approaches in this field, and their adaption to immunoprecipitation of AGO proteins allows for simultaneous identification of miRNAs along with mRNA targets and corresponding binding sites. In all three protocols, interactions between DNA and RNA-binding proteins and their respective targets are initially fixed by UV cross-linking. Following precipitation and purification of AGO(s) as the RNA-binding protein(s) of interest, cross-linked and co-immunoprecipitated RNAs are partially digested and subsequently released by proteinase digestion. The remaining RNA fragments are adaptor-ligated, reverse-transcribed into cDNA, amplified and finally subjected to next-generation sequencing (NGS). While HITS-CLIP and PAR-CLIP provide distinct sets of reads that are either derived from miRNAs or target mRNAs, CLASH involves an additional ligation step subsequent to partial digestion of protein-bound RNA and prior to adaptor ligation. This additional intramolecular ligation results in generation of chimeric reads that comprise a respective miRNA sequence along with the targeted mRNA. Using this approach, interaction sites of the RNA molecules that gave rise to a chimeric read can be identified by in silico folding. The localization of protein RNA-binding sites within sequencing data from CLIP experiments is based on identification of cross-linking induced mutation sites (CIMS). Reads generated by HITS-CLIP are screened for cross-linking induced deletions that originate during reverse transcription of RNAs with amino-acid-RNA adducts [90]. In contrast to substitutions or insertions, these deletions are highly site-specific and permit identification of interaction sites at single-nucleotide resolution. Likewise, reads generated by PAR-CLIP are screened for CIMS to identify binding regions. Since PAR-CLIP aims to maximize cross-linking efficiency by random incorporation of photoactivatable nucleoside analogues into nascent RNAs, CIMS are represented by specific substitutions within the sequenced reads. Feeding of cultured cells with 4-thiouridine, for instance, results in thymidine to cytidine transitions in the cross-linked RNA. Once the binding regions have been defined by target CIMS, miRNA–mRNA interactions can be deduced by comparison of corresponding regions between the miRNA and mRNA reads from HITS-CLIP or PAR-CLIP experiments.

Using PAR-CLIP, $\sim$3500 canonical miRNA target interactions could be identified in MCF-7 cells and the combined analysis of these with gene expression profiles from individual breast cancer patients revealed a significant correlation between expression levels of miR-182 targeted mRNAs and overall patient survival [91]. MiR-218 is significantly downregulated in human medulloblastoma and a recent screening for corresponding target sites via HITS-CLIP identified more than 600 target genes including both previously known and completely novel targets [92]. Since miRNA–mRNA interactions are directly preserved within chimeric reads, CLASH provides a quantitative basis for analysis of such interactions. The chimeric reads from human cell cultures confirmed that mRNAs represent the principal binding partner of miRNAs ($\sim$70 % of the identified interactions). The remaining binding partners comprised pseudogenes, long intergenic ncRNAs, ribosomal RNAs, transfer RNAs, small nuclear RNAs, as well as ceRNAs, such as the circular RNA sponge for miR-7 (ciRS-7). In addition to the quantitative aspects of CLASH, miRNA–mRNA base pairing patterns can be deduced from chimeric reads without prior knowledge of targets or binding modes.

In silico analysis of miRNA-mRNA interactions in animals

As previously outlined, experimental methods for identification of miRNA–mRNA interactions in animals often require highly sophisticated bioinformatical analyses and/or precedent knowledge of promising candidate interactions. Consequently, computational methods have to be constantly improved and modulated depending on the available experimental data. Current computational approaches for in silico prediction of miRNA targets are discussed in this section with special emphasis on methods that combine data from mRNA and miRNA quantification experiments.

In silico prediction methods and interaction databases

To date, several algorithms for in silico prediction of miRNAs and their target genes have been developed. Prominent tools comprise TargetScan [93], miRanda [94, 95], PITA [96] and PicTar [97], which rely on sequence complementarity of the 3′ UTRs from potential mRNA targets with the seed region of a given miRNA. Additionally, these algorithms account for the secondary structure of the miRNA and/or respective target sites (e.g. free energy of the miRNA–mRNA union: ΔG or costs to unfold the secondary structure of the target site $\Delta\Delta G$), the number of potential binding sites within one transcript, and if applicable, conservation of the miRNA and/or the target site across mammals (assuming that conservation increases the likelihood of a functional site) [36, 98, 99].

Table 1 Overview of commonly used miRNA target databases

Database	URL	Last update	MiRBase	#Species	Expression profiles	Validated targets	Predicted targets
miRTarBase	http://mirtarbase.mbc.nctu.edu.tw	2013	V20	18	T	T	
TarBase	http://microrna.gr/tarbase	2013	NA	21		T	
miRecords	http://mirecords.biolead.org	2013	NA	9		T	T
miRNAMap	http://mirnamap.mbc.nctu.edu.tw	2007	V9.2	12	T		T
TargetScanS	http://targetscan.org	2012	V17	5			T
microRNA.org	http://microrna.org	2010	V15	5			T
miRDB	http://mirdb.org	2012	V18	5			T
miRGator	http://genome.ewha.ac.kr/miRGator	2007	V10	2	T		T
PicTar	http://pictar.mdc-berlin.de	2007	V9.2	4	T		T
RNA22	http://cm.jefferson.edu/rna22v2.0	2012	V18	2			T
DIANA MicroT	http://microrna.gr/microT-CDS	2012	V18	16			T
MicroInspector	http://mirna.imbb.forth.gr	2012	V17	141			T
miRo	http://ferrolab.dmi.unict.it/miro	2009	V12	1	T		T

The name of each tool and its web address is listed along with the year of the last update. Additionally, the table comprises the respective miRBase version of the current release, the number of contained species as well as Booleans, which indicate if expression profiles of miRNAs are available and if the database also contains validated and/or predicted targets. NA indicates missing information of the respective database about the underlying miRBase version

To avoid computationally intensive re-calculation of target predictions, the results of several of the tools are also available from database resources (Table 1). These databases include all predicted interactions between known 3′ UTRs of selected organisms and known miRNAs from miRBase [9]. Nevertheless, most of them are not updated regularly and none of them takes into account that 3′ UTR length of one and the same transcript is highly variable between different biological conditions due to APA [100]. This heterogeneity in 3′ UTR length represents one of the major reasons for the relatively high false positive rate of available predictions in databases, which is estimated to reach up to 70 % [101], and consequently complicates reliable predictions for interactions in a given tissue.

The overlaps between prediction results of different tools/databases are very low and range from 5 to 70 %. For that reason, databases such as miRwalk [102] or miRo [103] only list interactions that are commonly identified by several tools to increase the likelihood of predicting true interactions. However, even those databases that apply a combination of algorithms and tools for prediction are not capable of taking into account specific biological characteristics of a given sample. More recently published database tools try to account for these by including co-expression information into the prediction algorithm. MirCoX [104] and miRConnect [105] take advantage of the fast growing number of miRNA and mRNA sequencing data in publicly available repositories, and use the data to calculate significant negative correlations between mRNA and miRNA expression values under various biological conditions such as disease or stress states.

Databases that summarize experimentally validated interactions comprise miRecords [106], that provides 2705 records of interactions between 644 miRNAs and 1901 target genes in nine animal species, as well as TarBase [107], the largest manually curated database, indexing more than 65,000 validated interactions in 21 species. The validated interactions summarise the results from reporter assays and high-throughput experiments. However, with ∼200,000 estimated interactions occurring only in humans [108], these databases remain a relatively limited resource for miRNA target identification.

Combining quantitative expression data with sequence-based predictions

The attempt to use quantitative mRNA expression profiles from appropriately designed experiments for improvement of in silico miRNA target predictions has led to the development of several computational models that improve sequence-based predictions with experimental data [109–112]. These models combine mRNA and miRNA expression profiles from a particular experiment with information about predicted miRNA–mRNA interactions. The underlying frameworks are based on finding significant negative correlations between the expression of a miRNA and its potential target gene in the present experimental setup to increase the likelihood of a true interaction (Fig. 5). Several web services like omiRas [113], CPSS [114] or ncPRo-seq [115] provide the possibility to quantify and compare miRNA expression from raw sequencing libraries of two biological conditions, such as healthy and diseased
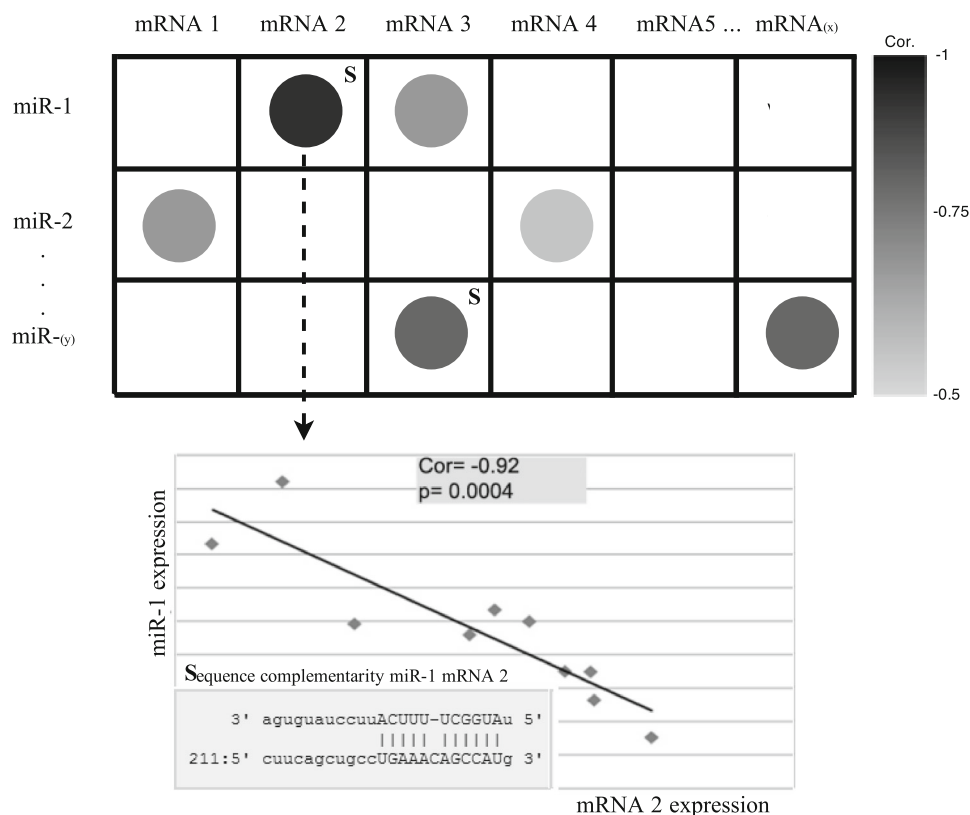
Fig. 5 Integration of experimental expression data into target prediction. In order to predict miRNA–mRNA interactions, a matrix for each miRNA-mRNA pair is generated (*top*). Potential miRNA–mRNA interactions exhibit a significant negative correlation (*circles*). Furthermore, interactions gain support by sequence complementarity (S), as visualized at the *left bottom*, and by the observed expression of miRNAs and mRNAs in each biological replicate (*dots* in the scatterplot at the *bottom*)

individuals. Furthermore, this quantitative data can be used to infer miRNA targets among differentially expressed genes between the same conditions, incorporating target predictions from various target databases. However, these models are solely focused on interactions with perfect seed matches resulting in post-transcriptional degradation of target mRNAs. Translational inhibition due to imperfect seed matches does not alter mRNA abundance and is therefore undetectable by correlation analysis. High-throughput measurement of protein levels could help overcome this limitation, but reliable and reproducible measurements of these is currently only achieved for a subset of proteins [116].

Including APA information into interaction predictions

Even though these approaches allow for considering the respective biological conditions under investigation, 3′ UTR heterogeneity remains disregarded. 3′ UTR shortening by APA can result in the absence of miRNA binding sites within a predicted target and consequently give rise to false positive predictions. Against this backdrop,

incorporation of information from 3′ end sequencing data into target prediction algorithms would significantly improve the prediction of interaction efficiency [117], and thus provide biologically more relevant results. To date, TargetScan Web service (http://targetscan.org/) represents the only miRNA–mRNA interaction database that includes according data into its prediction results. The predicted interactions for zebrafish miRNAs are available for each 3′ UTR isoform detected by 3P-Seq [58]. However, the abundance of specific isoforms is not taken into account to rank the interaction efficiencies. The upcoming version of TargetScan is going to include the fractions of mRNAs with a specific miRNA target site into the predictions for human, mouse and zebrafish, but these fractions will still not be tissue-specific.

Nam and colleagues were the first that introduced a revised prediction model, which weights prediction scores for a miRNA and its target site by the percentage of PA isoforms in each cell type that carry this interaction site [117]. On average, this model outperforms previous models by 50 %, thus indicating a new strategy in miRNA–mRNA interaction algorithm development. The growing number of global and tissue-specific APA profiles in databases such as APADB

(http://tools.genxpro.net/apadb/) provides the necessary resource for a significant improvement of miRNA–mRNA interaction predictions in diseased and stressed tissues.

Predicting the effect of SNPs on miRNA–mRNA interactions

Single-nucleotide polymorphisms can influence miRNA binding by two distinct mechanisms. On the one hand, SNPs can alter the seed region of a miRNA or mRNA binding site, which is linked to increased/decreased binding efficiencies. On the other, SNPs can interrupt or create novel PA signals that lead to transcript isoforms with altered miRNA binding site availability. A screening of human SNPs indicates that a substantial fraction of SNPs harbor the potential to create or disrupt APA signals, and this type of SNPs has been defined as APA-SNPs [118]. Many APA-SNPs are associated with shortened transcripts and increased gene expression due to reduced miRNA binding site availability, especially in disease. Moreover, these SNPs do not only affect PA site usage by altering canonical PA signals. An APA-SNP located in the U/GU-rich downstream element of the second PA signal of the human *ATP1B1* gene was shown to be associated with higher blood pressure in a European–American population [119]. This APA-SNP is linked to an increased usage of the second PA site due to increased CstF binding efficiency. The resulting transcript isoform exhibits increased translation efficiency and consequently yields a larger amount of protein product.

Recently, several databases that account for the impact of SNPs within 3′ UTRs on miRNA binding have been released [120–123]. Most of these rely on simple modifications of known algorithms like TargetScan or miRanda to detect the effects of a given SNP on miRNA binding stability [124]. However, these databases only rely on known SNPs from public resources like DBSNP and are therefore restricted to a predefined subset of SNPs [125]. To determine the influence of novel variations, mrSNP [126] calculates the impact of user supplied SNPs based on different criteria. Most importantly, these criteria comprise the distance in the alignment to the seed region of a miRNA, and the changes in the energy of the alignment compared to an alignment without the mutation. Currently, limited data for evaluation is available, but mrSNP correctly predicted disrupted binding for 69 % (11/16) of the SNPs with experimentally verified effect on miRNA binding.

Concluding remarks and perspectives

The functional fraction of the human genome remains a widely discussed topic and naturally depends on the definition of functionality. While the ENCODE sequencing project revealed that more than 80 % of the human genome codes for RNA, recent analysis that takes evolutionary conservation into account estimates 7.1–9 % of the human genome to be functional [127]. However, only a small fraction of these functional sites (1 % of the genome) is protein-coding. Comparative analyses between mouse and human tissues suggest that most, but not all, non-coding RNAs have short-lived lineage-specific functionality, which is especially true for long non-coding RNAs. In contrast, miRNA precursors and in particular, seed regions of mature miRNAs are highly conserved between mammals which underline their role as important post-transcriptional regulators.

The advent of NGS methods that allow for global investigation of the miRNAome across several tissues and disease states along with corresponding gene expression profiles provides the basis for a deeper understanding of the post-transcriptional regulation of biological pathways. Experimental methods such as CLASH permit a reliable detection and even quantification of direct miRNA–mRNA target interactions on a global scale. This trend is going to further extend the overwhelmingly large amount of sequencing data in respective databases, and consequently more efficient algorithms and data-mining tools are needed. The resulting gain of knowledge will likely improve therapeutic options for treatments of serious diseases such as cancer. In 2013, the first clinical study based on a miRNA mimic was initiated and more miRNAs, currently published as biomarkers with unknown function, are waiting in line for potential clinical use. To accelerate this process, the underlying molecular mechanisms of these miRNAs, and especially their condition and tissue-specific target genes, need further investigation. APA is one of the major reasons for 3′ UTR heterogeneity, and consequently, identification of condition and tissue-specific PA isoforms represents the next necessary step to fill this gap in knowledge. With the growing number of 3′ end sequencing data that is available in public databases, regulatory models can be further improved not only in the context of human diseases, but also for other animals and cellular processes.

References

1. Carthew RW, Sontheimer EJ (2009) Origins and mechanisms of miRNAs and siRNAs. Cell 136(4):642–655
2. Cullen BR (2006) Viruses and microRNAs. Nature Genet 38:S25–S30
3. Huang PJ, Lin WC, Chen SC, Lin YH, Sun CH, Lyu PC, Tang P (2012) Identification of putative miRNAs from the deep-branching unicellular flagellates. Genomics 99(2):101–107

4. Lee HC, Li L, Gu W, Xue Z, Crosthwaite SK, Pertsemlidis A, Lewis ZA, Freitag M, Selker EU, Mello CC, Liu Y (2010) Diverse pathways generate microRNA-like RNAs and Dicer-independent small interfering RNAs in fungi. Mol Cell 38(6):803–814

5. Molnar A, Schwach F, Studholme DJ, Thuenemann EC, Baulcombe DC (2007) miRNAs control gene expression in the single-cell alga Chlamydomonas reinhardtii. Nature 447(7148):1126–1129

6. Yao Y, Nair V (2014) Role of virus-encoded microRNAs in Avian viral diseases. Viruses 6(3):1379–1394

7. 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

8. Kozomara A, Griffiths-Jones S (2013) miRBase: annotating high confidence microRNAs using deep sequencing data. Nucleic Acids Res. doi:10.1093/nar/gkt1181

9. Kozomara A, Griffiths-Jones S (2010) miRBase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res. doi:10.1093/nar/gkq1027

10. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ (2006) miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 34(Database issue):D140–D144

11. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ (2008) miRBase: tools for microRNA genomics. Nucleic Acids Res 36(Database issue):D154–D158

12. Bazzini AA, Lee MT, Giraldez AJ (2012) Ribosome profiling shows that miR-430 reduces translation before causing mRNA decay in zebrafish. Science 336(6078):233–237

13. Bethune J, Artus-Revel CG, Filipowicz W (2012) Kinetic analysis reveals successive steps leading to miRNA-mediated silencing in mammalian cells. EMBO Rep 13(8):716–723

14. Djuranovic S, Nahvi A, Green R (2012) miRNA-mediated gene silencing by translational repression followed by mRNA deadenylation and decay. Science 336(6078):237–240

15. Vasudevan S, Tong Y, Steitz JA (2007) Switching from repression to activation: microRNAs can up-regulate translation. Science 318(5858):1931–1934

16. Morlando M, Ballarino M, Gromak N, Pagano F, Bozzoni I, Proudfoot NJ (2008) Primary microRNA transcripts are processed co-transcriptionally. Nat Struct Mol Biol 15(9):902–909

17. Fahlgren N, Howell MD, Kasschau KD, Chapman EJ, Sullivan CM, Cumbie JS, Givan SA, Law TF, Grant SR, Dangl JL, Carrington JC (2007) High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of MIRNA genes. PLoS One 2(2):e219

18. Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ (2004) Processing of primary microRNAs by the Microprocessor complex. Nature 432(7014):231–235

19. Kim VN (2005) MicroRNA biogenesis: coordinated cropping and dicing. Nat Rev Mol Cell Biol 6(5):376–385

20. Yi R, Qin Y, Macara IG, Cullen BR (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev 17(24):3011–3016

21. Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH (2001) Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in C. elegans. Genes Dev 15(20):2654–2659

22. Lee YS, Nakahara K, Pham JW, Kim K, He Z, Sontheimer EJ, Carthew RW (2004) Distinct roles for Drosophila Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. Cell 117(1):69–81

23. Czech B, Hannon GJ (2011) Small RNA sorting: matchmaking for Argonautes. Nat Rev Genet 12(1):19–31

24. Voinnet O (2009) Origin, biogenesis, and activity of plant microRNAs. Cell 136(4):669–687

25. Ladewig E, Okamura K, Flynt AS, Westholm JO, Lai EC (2012) Discovery of hundreds of mirtrons in mouse and human small RNA data. Genome Res 22(9):1634–1645

26. Ruby JG, Jan CH, Bartel DP (2007) Intronic microRNA precursors that bypass Drosha processing. Nature 448(7149):83–86

27. Meng Y, Shao C (2012) Large-scale identification of mirtrons in Arabidopsis and rice. PLoS One 7(2):e31163

28. Zhu QH, Spriggs A, Matthew L, Fan L, Kennedy G, Gubler F, Helliwell C (2008) A diverse set of microRNAs and microRNA-like small RNAs in developing rice grains. Genome Res 18(9):1456–1465

29. Hertel J, Langenberger D, Stadler PF (2014) Computational prediction of microRNA genes. Methods Mol Biol 1097:437–456

30. Liu J, Carmell MA, Rivas FV, Marsden CG, Thomson JM, Song JJ, Hammond SM, Joshua-Tor L, Hannon GJ (2004) Argonaute2 is the catalytic engine of mammalian RNAi. Science 305(5689):1437–1441

31. Song JJ, Smith SK, Hannon GJ, Joshua-Tor L (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. Science 305(5689):1434–1437

32. Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, Voinnet O (2008) Widespread translational inhibition by plant miRNAs and siRNAs. Science 320(5880):1185–1190

33. Bartel DP (2009) MicroRNAs: target recognition and regulatory functions. Cell 136(2):215–233

34. Axtell MJ, Westholm JO, Lai EC (2011) Vive la difference: biogenesis and evolution of microRNAs in plants and animals. Genome Biol 12(4):221

35. Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP (2002) Prediction of plant microRNA targets. Cell 110(4):513–520

36. Grimson A, Farh KK-H, 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

37. Stroynowska-Czerwinska A, Fiszer A, Krzyzosiak WJ (2014) The panorama of miRNA-mediated mechanisms in mammalian cells. Cell Mol Life Sci 71(12):2253–2270

38. Nicoloso MS, Sun H, Spizzo R, Kim H, Wickramasinghe P, Shimizu M, Wojcik SE, Ferdin J, Kunej T, Xiao L et al (2010) Single-nucleotide polymorphisms inside microRNA target sites influence tumor susceptibility. Cancer Res 70(7):2789–2798

39. Chen K, Song F, Calin GA, Wei Q, Hao X, Zhang W (2008) Polymorphisms in microRNA targets: a gold mine for molecular epidemiology. Carcinogenesis 29(7):1306–1311

40. Elkon R, Ugalde AP, Agami R (2013) Alternative cleavage and polyadenylation: extent, regulation and function. Nat Rev Genet 14(7):496–506

41. Tian B, Manley JL (2013) Alternative cleavage and polyadenylation: the long and short of it. Trends Biochem Sci 38(6):312–320

42. Sandberg R, Neilson JR, Sarma A, Sharp PA, Burge CB (2008) Proliferating cells express mRNAs with shortened 3' untranslated regions and fewer microRNA target sites. Science 320(5883):1643–1647

43. Millevoi S, Vagner S (2010) Molecular mechanisms of eukaryotic pre-mRNA 3′ end processing regulation. Nucleic Acids Res 38(9):2757–2774

44. Proudfoot NJ (2011) Ending the message: poly(A) signals then and now. Genes Dev 25(17):1770–1782

45. Hu J, Lutz CS, Wilusz J, Tian B (2005) Bioinformatic identification of candidate cis-regulatory elements involved in human mRNA polyadenylation. RNA 11(10):1485–1493

46. Scorilas A (2002) Polyadenylate polymerase (PAP) and 3′ end pre-mRNA processing: function, assays, and association with disease. Crit Rev Clin Lab Sci 39(3):193–224

47. Di Giammartino DC, Manley JL (2014) New links between mRNA polyadenylation and diverse nuclear pathways. Mol Cells 37(9):644–649

48. Shi Y, Di Giammartino DC, Taylor D, Sarkeshik A, Rice WJ, Yates JR 3rd, Frank J, Manley JL (2009) Molecular architecture of the human pre-mRNA 3' processing complex. Mol Cell 33(3):365–376

49. Elkon R, Drost J, van Haaften G, Jenal M, Schrier M, Oude Vrielink JA, Agami R (2012) E2F mediates enhanced alternative polyadenylation in proliferation. Genome Biol 13(7):R59

50. Han T, Kim JK (2014) Driving glioblastoma growth by alternative polyadenylation. Cell Res 24(9):1023–1024

51. Miura P, Shenker S, Andreu-Agullo C, Westholm JO, Lai EC (2013) Widespread and extensive lengthening of 3' UTRs in the mammalian brain. Genome Res 23(5):812–825

52. Di Giammartino DC, Nishida K, Manley JL (2011) Mechanisms and consequences of alternative polyadenylation. Mol Cell 43(6):853–866

53. Kaida D, Berg MG, Younis I, Kasim M, Singh LN, Wan L, Dreyfuss G (2010) U1 snRNP protects pre-mRNAs from premature cleavage and polyadenylation. Nature 468(7324):664–668

54. Berg MG, Singh LN, Younis I, Liu Q, Pinto AM, Kaida D, Zhang Z, Cho S, Sherrill-Mix S, Wan L et al (2012) U1 snRNP determines mRNA length and regulates isoform expression. Cell 150(1):53–64

55. Huang H, Chen J, Liu H, Sun X (2013) The nucleosome regulates the usage of polyadenylation sites in the human genome. BMC genomics 14912

56. Shepard PJ, Choi EA, Lu J, Flanagan LA, Hertel KJ, Shi Y (2011) Complex and dynamic landscape of RNA polyadenylation revealed by PAS-Seq. RNA 17(4):761–772

57. Derti A, Garrett-Engele P, Macisaac KD, Stevens RC, Sriram S, Chen R, Rohl CA, Johnson JM, Babak T (2012) A quantitative atlas of polyadenylation in five mammals. Genome Res 22(6):1173–1183

58. Jan CH, Friedman RC, Ruby JG, Bartel DP (2011) Formation, regulation and evolution of Caenorhabditis elegans 3 [prime] UTRs. Nature 469(7328):97–101

59. Zawada AM, Rogacev KS, Muller S, Rotter B, Winter P, Fliser D, Heine GH (2014) Massive analysis of cDNA Ends (MACE) and miRNA expression profiling identifies proatherogenic pathways in chronic kidney disease. Epigenetics 9(1):161–172

60. Hoque M, Ji Z, Zheng D, Luo W, Li W, You B, Park JY, Yehia G, Tian B (2013) Analysis of alternative cleavage and polyadenylation by 3' region extraction and deep sequencing. Nat Methods 10(2):133–139

61. Müller S, Rycak L, Afonso-Grunz F, Winter P, Zawada AM, Damrath E, Scheider J, Schmäh J, Koch I, Kahl G, others (2014) APADB: a database for alternative polyadenylation and microRNA regulation events. Database. doi:10.1093/database/bau076

62. Leslie C (2014) Context-specific 3'UTR isoform expression and miRNA regulation. In Intergrative RNA Biology Special Interest Group Meeting, p 16

63. Ulitsky I, Shkumatava A, Jan CH, Subtelny AO, Koppstein D, Bell GW, Sive H, Bartel DP (2012) Extensive alternative polyadenylation during zebrafish development. Genome Res 22(10):2054–2066

64. Ji Z, Lee JY, Pan Z, Jiang B, Tian B (2009) Progressive lengthening of 3' untranslated regions of mRNAs by alternative polyadenylation during mouse embryonic development. Proc Natl Acad Sci USA 106(17):7028–7033

65. Ji Z, Tian B (2009) Reprogramming of 3' untranslated regions of mRNAs by alternative polyadenylation in generation of pluripotent stem cells from different cell types. PLoS One 4(12):e8419

66. Müller S (2014) In silico analysis of regulatory networks underlines the role of miR-10b-5p and its target BDNF in huntington's disease. Transl Neurodegr 3(1):17

67. Zuccato C, Cattaneo E (2007) Role of brain-derived neurotrophic factor in Huntington's disease. Prog Neurobiol 81(5–6):294–330

68. Varendi K, Kumar A, Härma M-A, Andressoo JO (2014) miR-1, miR-10b, miR-155, and miR-191 are novel regulators of BDNF. Cell Mol Life Sci 71(22):4443–4456. doi:10.1007/s00018-014-1628-x

69. 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

70. Lee RC, Feinbaum RL, Ambros V (1993) The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 75(5):843–854

71. Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, Horvitz HR, Ruvkun G (2000) The 21-nucleotide let-7 RNA regulates developmental timing in Caenorhabditis elegans. Nature 403(6772):901–906

72. Wightman B, Ha I, Ruvkun G (1993) Posttranscriptional regulation of the heterochronic gene lin-14 by lin-4 mediates temporal pattern formation in C. elegans. Cell 75(5):855–862

73. Kuhn DE, Martin MM, Feldman DS, Terry AV Jr, Nuovo GJ, Elton TS (2008) Experimental validation of miRNA targets. Methods 44(1):47–54

74. German MA, Pillay M, Jeong DH, Hetawal A, Luo S, Janardhanan P, Kannan V, Rymarquis LA, Nobuta K, German R, De Paoli E, Lu C, Schroth G, Meyers BC, Green PJ (2008) Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. Nat Biotechnol 26(8):941–946

75. Gregory BD, O'Malley RC, Lister R, Urich MA, Tonti-Filippini J, Chen H, Millar AH, Ecker JR (2008) A link between RNA metabolism and silencing affecting Arabidopsis development. Dev Cell 14(6):854–866

76. Addo-Quaye C, Eshoo TW, Bartel DP, Axtell MJ (2008) Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome. Curr Biol 18(10):758–762

77. Karlova R, van Haarst JC, Maliepaard C, van de Geest H, Bovy AG, Lammers M, Angenent GC, de Maagd RA (2013) Identification of microRNA targets in tomato fruit development using high-throughput sequencing and degradome analysis. J Exp Bot 64(7):1863–1878

78. Shamimuzzaman M, Vodkin L (2012) Identification of soybean seed developmental stage-specific and tissue-specific miRNA targets by degradome sequencing. BMC Genom 13:310

79. Addo-Quaye C, Miller W, Axtell MJ (2009) CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. Bioinformatics 25(1):130–131

80. Folkes L, Moxon S, Woolfenden HC, Stocks MB, Szittya G, Dalmay T, Moulton V (2012) PAREsnip: a tool for rapid genome-wide discovery of small RNA/target interactions evidenced through degradome sequencing. Nucleic Acids Res 40(13):e103

81. Willmann MR, Berkowitz ND, Gregory BD (2014) Improved genome-wide mapping of uncapped and cleaved transcripts in eukaryotes–GMUCT 2.0. Methods 67(1):64–73

82. Zhai J, Arikit S, Simon SA, Kingham BF, Meyers BC (2014) Rapid construction of parallel analysis of RNA end (PARE) libraries for Illumina sequencing. Methods 67(1):84–90

83. Bader AG, Brown D, Winkler M (2010) The promise of microRNA replacement therapy. Cancer Res 70(18):7027–7030

84. Krützfeldt J, Rajewsky N, Braich R, Rajeev KG, Tuschl T, Manoharan M, Stoffel M (2005) Silencing of microRNAs in vivo with 'antagomirs'. Nature 438(7068):685–689

85. Thomas M, Lieberman J, Lal A (2010) Desperately seeking microRNA targets. Nat Struct Mol Biol 17(10):1169–1174

86. Thomson DW, Bracken CP, Goodall GJ (2011) Experimental strategies for microRNA target identification. Nucleic Acids Res 39(16):6845–6853

87. Chi SW, Zang JB, Mele A, Darnell RB (2009) Argonaute HITS-CLIP decodes microRNA-mRNA interaction maps. Nature 460(7254):479–486

88. Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, Rothballer A, Ascano M Jr, Jungkamp A-C, Munschauer M et al (2010) Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 141(1):129–141

89. 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

90. Zhang C, Darnell RB (2011) Mapping in vivo protein-RNA interactions at single-nucleotide resolution from HITS-CLIP data. Nat Biotechnol 29(7):607–614

91. Farazi TA, Hoeve J, Brown M, Mihailovic A, Horlings HM, Vijver MVD, Tuschl T, Wessels L (2014) Identification of distinct miRNA target regulation between breast cancer molecular subtypes using AGO2-PAR-CLIP and patient datasets. Genome Biol 15:R9

92. Venkataraman S, Birks DK, Balakrishnan I, Alimova I, Harris PS, Patel PR, Handler MH, Dubuc A, Taylor MD, Foreman NK et al (2013) MicroRNA 218 acts as a tumor suppressor by targeting multiple cancer phenotype-associated genes in medulloblastoma. J Biol Chem 288(3):1918–1928

93. 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(1):15–20

94. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS et al (2004) MicroRNA targets in Drosophila. Genome Biol 5(1):R1

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

96. 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

97. Krek A, Grün D, Poy MN, Wolf R, Rosenberg L, Epstein EJ, MacMenamin P, da Piedade I, Gunsalus KC, Stoffel M et al (2005) Combinatorial microRNA target predictions. Nat Genet 37(5):495–500

98. Friedman RC, Farh KK-H, Burge CB, Bartel DP (2009) Most mammalian mRNAs are conserved targets of microRNAs. Genome Res 19(1):92–105

99. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB et al (2003) Prediction of mammalian microRNA targets. Cell 115(7):787–798

100. Lee D, Shin C (2012) MicroRNA-target interactions: new insights from genome-wide approaches. Ann N Y Acad Sci 1271(1):118–128

101. Muniategui A, Pey J, Planes FJ, Rubio A (2013) Joint analysis of miRNA and mRNA expression data. Brief Bioinform 14(3):263–278

102. Dweep H, Sticht C, Pandey P, Gretz N (2011) miRWalk-database: prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform 44(5):839–847

103. Laganà A, Forte S, Giudice A, Arena M, Puglisi P, Giugno R, Pulvirenti A, Shasha D, Ferro A (2009) miRò: a miRNA knowledge base. Database 2009:bap008

104. Giles CB, Girija-Devi R, Dozmorov MG, Wren JD (2013) mirCoX: a database of miRNA-mRNA expression correlations derived from RNA-seq meta-analysis. BMC Bioinform 14(Suppl 14):S17

105. Hua Y, Duan S, Murmann AE, Larsen N, Kjems J, Lund AH, Peter ME (2011) miRConnect: identifying effector genes of

106. 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(suppl 1):D105–D110

107. 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(D1):D222–D229

108. Brennecke J, Stark A, Russell RB, Cohen SM (2005) Principles of microRNA-target recognition. PLoS Biol 3(3):e85

109. Carmona-Saez P, Chagoyen M, Tirado F, Carazo JM, Pascual-Montano A (2007) GENECODIS: a web-based tool for finding significant concurrent annotations in gene lists. Genome Biol 8(1):R3

110. Li J, Min R, Bonner A, Zhang Z (2009) A probabilistic framework to improve microrna target prediction by incorporating proteomics data. J Bioinform Comput Biol 7(06):955–972

111. Nam S, Li M, Choi K, Balch C, Kim S, Nephew KP (2009) MicroRNA and mRNA integrated analysis (MMIA): a web tool for examining biological functions of microRNA expression. Nucleic Acids Res 37(suppl 2):W356–W362

112. Ritchie W, Rajasekhar M, Flamant S, Rasko JE (2009) Conserved expression patterns predict microRNA targets. PLoS Comput Biol 5(9):e1000513

113. Muller S, Rycak L, Winter P, Kahl G, Koch I, Rotter B (2013) omiRas: a Web server for differential expression analysis of miRNAs derived from small RNA-Seq data. Bioinformatics 29(20):2651–2652

114. Zhang Y, Xu B, Yang Y, Ban R, Zhang H, Jiang X, Cooke HJ, Xue Y, Shi Q (2012) CPSS: a computational platform for the analysis of small RNA deep sequencing data. Bioinformatics 28(14):1925–1927

115. Chen CJ, Servant N, Toedling J, Sarazin A, Marchais A, Duvernois-Berthet E, Cognat V, Colot V, Voinnet O, Heard E, Ciaudo C, Barillot E (2012) ncPRO-seq: a tool for annotation and profiling of ncRNAs in sRNA-seq data. Bioinformatics 28(23):3147–3149

116. Doerr A (2013) Mass spectrometry-based targeted proteomics. Nat Methods 10(1):23

117. 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

118. Thomas LF, S\aetrom Pa (2012) Single nucleotide polymorphisms can create alternative polyadenylation signals and affect gene expression through loss of microRNA-regulation. PLoS Comput Biol 8(8):e1002621

119. Prasad MK, Bhalla K, Pan ZH, O'Connell JR, Weder AB, Chakravarti A, Tian B, Chang YP (2013) A polymorphic 3′UTR element in ATP1B1 regulates alternative polyadenylation and is associated with blood pressure. PLoS One 8(10):e76290

120. Barenboim M, Zoltick BJ, Guo Y, Weinberger DR (2010) MicroSNiPer: a web tool for prediction of SNP effects on putative microRNA targets. Hum Mutat 31(11):1223–1232

121. 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 Genom 13(1):44

122. Hiard S, Charlier C, Coppieters W, Georges M, Baurain D (2010) Patrocles: a database of polymorphic miRNA-mediated gene regulation in vertebrates. Nucleic Acids Res 38(suppl 1):D640–D651

123. Ziebarth JD, Bhattacharya A, Chen A, Cui Y (2011) PolymiRTS Database 2.0: linking polymorphisms in microRNA target sites

miRNAs and miRNA families in cancer cells. PLoS One 6(10):e26521

🕑 Springer

with human diseases and complex traits. Nucleic Acids Res. doi:10.1093/nar/gkr1026

124. Liu C, Zhang F, Li T, Lu M, Wang L, Yue W, Zhang D (2012) MirSNP, a database of polymorphisms altering miRNA target sites, identifies miRNA-related SNPs in GWAS SNPs and eQTLs. BMC Genom 13(1):661

125. Sherry ST, Ward M, Sirotkin K (1999) dbSNP-database for single nucleotide polymorphisms and other classes of minor genetic variation. Genome Res 9(8):677–679

126. Deveci M, Catalyürek ÜV, Toland AE (2014) mrSNP: software to detect SNP effects on microRNA binding. BMC Bioinform 15(1):73

127. Rands CM, Meader S, Ponting CP, Lunter G (2014) 8.2% of the Human genome is constrained: variation in rates of turnover across functional element classes in the human lineage. PLoS Genet 10(7):e1004525