Review

Finding MicroRNA Targets in Plants: Current Status and Perspectives

Jiandong Ding [1,2,†], Shuigeng Zhou [1,2,*], Jihong Guan [3,*]

[1] Shanghai Key Laboratory of Intelligent Information Processing, Fudan University, Shanghai 200433, China
[2] School of Computer Science, Fudan University, Shanghai 200433, China
[3] Department of Computer Science and Technology, Tongji University, Shanghai 201804, China

Abstract

MicroRNAs (miRNAs), a class of ∼20–24 nt long non-coding RNAs, have critical roles in diverse biological processes including development, proliferation, stress response, *etc*. With the development and availability of experimental technologies and computational approaches, the field of miRNA biology has advanced tremendously over the last decade. By sequence complementarity, miRNAs have been estimated to regulate certain mRNA transcripts. Although it was once thought to be simple and straightforward to find plant miRNA targets, this viewpoint is being challenged by genetic and biochemical studies. In this review, we summarize recent progress in plant miRNA target recognition mechanisms, principles of target prediction, and introduce current experimental and computational tools for plant miRNA target prediction. At the end, we also present our thinking on the outlook for future directions in the development of plant miRNA target finding methods.

Keywords: MicroRNA; Target prediction; Degradome-seq; Integration

Introduction

MicroRNAs (miRNAs) are one of three distinct types of small RNAs (sRNAs, including small interfering RNAs, microRNAs and Piwi-interacting RNAs) currently understood in plants and animals, which are distinguished by their biogenesis, not by their action [1]. The field of miRNA biology emerged with the discovery that the gene *lin-4* and *let-7*, which control developmental timing in the nematode *Caenorhabditis elegans*, surprisingly did not code for protein, but instead acted as a ∼22 nt RNA transcript which regulated gene expression post-transcriptionally [2–4]. Since then, miRNAs have been shown to play a variety of regulatory roles and target other regions in addition to 3′ untranslated regions (UTRs) [5–8]. The field of miRNA biology has quickly progressed through the application of genome-wide approaches for identification of

miRNAs [9,10] and their targets [11–14]. In this article, we focus on plants and review these tools and their contributions to our understanding of miRNAs as local and global regulators.

The biogenesis of plant miRNAs has been documented most in *Arabidopsis thaliana* (**Figure 1**). Mature plant miRNAs range in size from 20 to 24 nucleotides. Primary miRNA transcripts (pri-miRNAs) are generally RNA polymerase II transcripts that contain imperfect, self-complementary foldback regions [15,16]. The length of plant pri-miRNA hairpins is heterogeneous, ranging from approximately 70 to thousands of bases.

In animals, the pri-miRNA transcript is first processed by the RNase III domain-containing protein *Drosha* in association with the RNA-binding protein encoded by DGCR8 [17]. Processed miRNA precursors (pre-miRNA) are exported from the nucleus and are cleaved by 22 bp from the *Drosha* processing site by the RNase III domain-containing protein *Dicer* [18]. In plants, there is no *Drosha* homolog present; rather, plants have Dicer homologs. Of the four Dicer-like (DCL) enzymes in

* Corresponding authors.
E-mail: sgzhou@fudan.edu.cn (Zhou S), jhguan@tongji.edu.cn (Guan J).
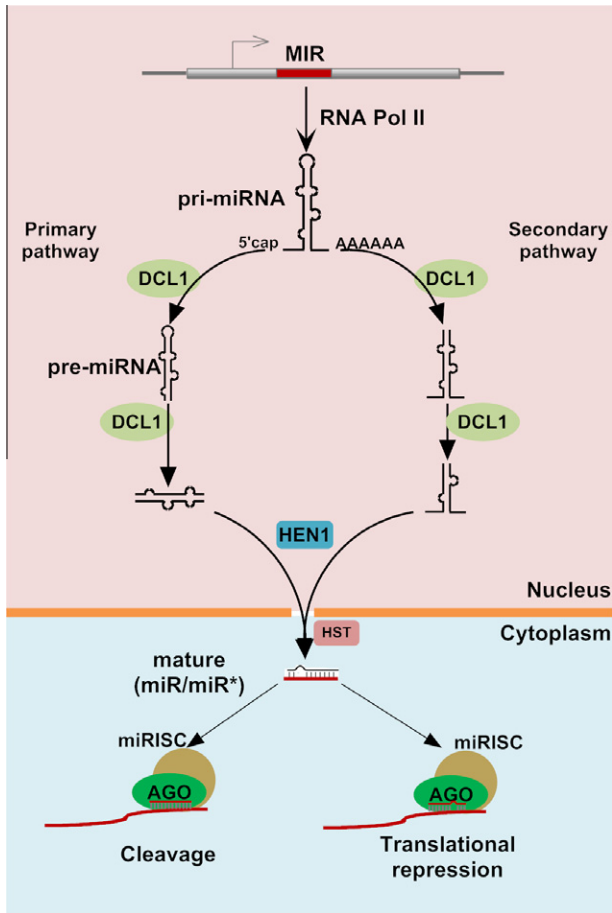† Current address: China Research Lab, IBM, Shanghai 201203, China.

Figure 1   Major biogenesis pathways of plant miRNAs

The figure shows two major pathways for plant miRNA biogenesis. The pri-miRNA is a primary transcript. The stem–loops on pri-miRNA are cleaved by DCL1 (Dicer-LIKE 1) in the nucleus giving rise to the mature transcript. The generated duplex is shown with a red strand (the miRNA) and a black strand (the miRNA*). Before exported into the cytosol by HASTY (HST), it is methylated by HUA ENHANCER 1 (HEN1) to render stability. The red strand is integrated in the miRISC and the black strand is either degraded or acts like the red strand. Depending on the degree of complementarity to the target site, miRISC will either cleave the mRNA which will induce immediate degradation or suppress translation procedure.

Arabidopsis, DCL1 is responsible for the bulk of miRNA biogenesis [19]. There are two precursor-processing pathways that have been identified for plant miRNA genes. The primary pathway involves stem-to-loop processing in which the sequence and structure beyond the miRNA–miRNA* site are necessary and used by the cleavage pathway components to excise the mature sequences [20–22]. The second pathway involves loop-to-stem processing in which only the structure between the miRNA and miRNA* is necessary for the cleavage pathway components to excise the mature sequences [23,24]. Most plant pri-miRNA hairpins produce a single miRNA/miRNA* duplex, but some loci consistently produce multiple duplexes [25].

Unlike animals, miRNA biogenesis in plants is completed within the nucleus [26]. Several accessory factors also contribute to the efficiency and fidelity of miRNA/miRNA* excision in plants [27,28]. The 3′ end of the initial

miRNA/miRNA* duplex is 2′-*O*-methylated by the nuclear HEN1 protein [29]; this modification prevents non-templated 3′-polymerization that accelerates miRNA turnover [30]. HASTY (HST), a plant homolog of Exportin-5, is then thought to export miRNA/miRNA* duplexes for loading into ARGONAUTE (AGO) proteins [31]. Usually, AGO1 acts as a 'slicer' to direct the endonucleolytic cleavage of target RNAs [32], although most other plant AGOs are also likely to possess slicing capabilities [33].

Many animal miRNAs are conserved even between greatly-diverged species [34], especially when considering the crucial 'seed' regions, which are often solely responsible for targeting specificity [35]. In plants, a minority of annotated MIR gene families are conserved between plant families, while the majority are family- or species-specific, suggesting that most known MIR genes arose relatively recently in evolutionary time [36]. Unlike highly-conserved, ancient miRNAs, young miRNAs are often weakly expressed and processed imprecisely, lack targets, and display patterns of neutral variation, suggesting that young MIR loci tend to evolve neutrally (for a recent review, see [37]). However, the evidence for any miRNAs conserved between animals and plants is slim, which suggests that animal and plant miRNAs may not come from a same 'ancestor'.

After combining the AGO protein, an outstanding problem in the miRNA field is how miRNAs recognize specific sequences of partial complementarity, complicating the prediction of target sites [38].

The mechanism by which miRNAs regulate target gene expression has been a controversial subject, as there is evidence for target mRNA destabilization, translational repression and even activation of gene expression [39]. The substantial differences between the biogenesis of animal and plant miRNAs are also reflected in the differences in their requirements for target recognition. In plants, miRNAs can silence targets through RNA degradation as well as translational repression pathways [40]. The perfect [41], or near perfect pairing of miRNA and its target site supports endonucleolytic cleavage of the mRNA by AGO (Figure 1). This mechanism is common in plants but much rarer in animals [41]. There are also cases in plants in which miRNAs cause reduced levels of protein, but not mRNA, suggesting that translational repression is directed by miRNA-induced silencing complex (miRISC) [42]. The actual mechanism that blocks protein production is not clear and there is evidence for inhibition of translational initiation or elongation, as well as for directed proteolysis of the peptide that is being synthesized from the targeted mRNA [43].

Overall, miRNAs typically repress gene expression, and it remains to be seen whether positive regulation of targets extends beyond the limited cases that have been uncovered so far [44]. Methods for discerning these different mechanisms of target regulation will be discussed in the next section.

In addition to the repression of target mRNAs, some miRNAs have other specialized functions or confer unique

properties to miRNA–AGO complexes. For example, some miRNAs can trigger the production of 21-nt siRNAs. *Trans*-acting siRNAs (ta-siRNAs) are RDR6- and DCL4-dependent products of a refined RNA interference pathway, which function as repressors on specific, co-evolved target mRNAs [45,46]. Phasing of TAS1/TAS2, TAS3 and TAS4 ta-siRNAs is set by cleavage guided by miR173–AGO1, miR390–AGO7 and miR828–AGO1 complexes, respectively [47–49]. Thus, miR173, miR390 and miR828 play as activators, rather than repressors, of a siRNA pathway. miRNAs can also direct DNA methylation. A subset of miRNA variants preferentially associate with AGO proteins involved in RNA-directed DNA methylation (RdDM) [50]. In rice, DCL3a also processes multiple MIR foldbacks, yielding 24-nt siRNA-like miRNAs [51]. Unlike 21-nt miRNAs, the 24-nt siRNA-like miRNAs preferentially associate with rice AGO4a and AGO4b and guide the methylation of target genes.

In order to facilitate the computational study of miR-NAs, especially the miRNA−target interactions, we summarized common online sources and grouped them into several categories (**Table 1**).

Generally, sequences of miRNA and transcript are required when studying miRNA−target interactions. As the biggest online registry, miRBase collects miRNAs of various species that are identified by experimental or computational methods [52]. Another well-known RNA database, Rfam, also provides miRNA sequences based on homology relationships [53]. Rfam is helpful both for miR-NA identification and for target prediction, especially when we need to filter out structural RNAs before analyzing deep sequencing data. Besides miRBase and Rfam, there are three plant-oriented miRNA annotation databases, MicroPC [54], PmiRKB [55] and PMRD [56], which greatly enhance plant miRNA studies. Several datasets are established to maintain plant genomic and/or transcriptomic data. Right now, TAIR [57] and TIGR [58] are such data centers for two well-studied plant model species *Arabidopsis* and rice, respectively. Additionally, Phytozome [59], incorporating information about 31 plant genomes and also their annotations, will be preferred when studying the evolution of plant species.

When estimating the performance of target prediction methods, it would be wise to use those experimentally-validated targets as test data. The emergence of miRTarBase [60] and starBase [61] satisfies such a requirement. Different from miRTarBase, starBase further integrates plant miR-NA−target interactions supported by Degradome-Seq. Along with the population of next generation sequencing [62], miRNA study is progressing rapidly. The appearance of MPSS [63], SRA [64] and GEO [65] sheds lights on the development of high-throughput based computational tools to study miRNAs in both animals and plants. Two other specific databases, ASRP [66] and CSRDB [67], are also listed in Table 1 because they perform a series of in-depth research related to sRNAs, including miRNAs.

Experimental methods

Target-specific validation methods

Given the challenge of matching miRNAs to specific target sequences, several approaches have been adopted for identifying functional interactions. In order to explore the principles of miRNA targeting mRNA, target-specific methods were introduced first. Target-specific experimental validation with well-established techniques such as quantitative real-time PCR (qRT-PCR) [68], western blot [69] and 5′-rapid amplification of cDNA ends (5′-RACE) [70], is commonly used to evaluate individual miRNA:target pairs. For a detailed review of methods for the experimental validation of specific miRNA targets, please refer to Thomson et al. [71].

Clearly the miRNA and its target should be co-expressed in order for the miRNA to regulate the expression of its biological target. Co-expression is typically demonstrated by simply performing Northern blot analysis or qRT-PCR using total RNA isolated from a specific cell type and probes or primers specific for a given miRNA and mRNA target [72].

Although the majority of miRNA targets appear to be regulated both at the mRNA and protein level, such regulation is only manifested at the protein level for some targets [73]. For a true target of a specific miRNA, the modulation of miRNA concentration should correspond to a predictable change in the amount of protein encoded by the target mRNA. Therefore, a typical approach to validate the functional importance of a miRNA:target pair is a transient over-expression of a given miRNA mimic in a cell type known to express the putative target protein and subsequent western analysis using a specific antibody against that protein. One limitation of monitoring protein concentration is that it may not be selective enough to distinguish between members of the same miRNA family with similar sequences.

Generally, the downstream effects of differential miR-NAs can be observed at the protein level by western blot and at the mRNA level by qRT-PCR, although these measures will not distinguish between direct and secondary miRNA targets and it is hard to determine whether the target mRNA is regulated by one single miRNA predominantly or several miRNAs simultaneously.

In the specific situation where a miRNA target is directly cleaved, 5′ RNA ligase mediated-RACE (5′ RLM-RACE) may be used to evaluate such targeting [41]. Briefly speaking, 5′-RACE is a PCR-based technique, whereby an RNA adapter is ligated to the free 5′ phosphate of an uncapped mRNA produced from, among other nucleolytic activities, AGO2-directed mRNA cleavage. The ligation product can be reversely transcribed using a forward primer directed against the linker and a gene specific reverse primer which is subsequently PCR amplified, cloned and identified by sequencing. 5′-RACE has been

Table 1   Online sources for plant miRNA study

Category	Name	miR	Target		Website	Note	Refs.
			Que[1]	Pre[2]			
miRNA databases	miRBase	√			http://www.mirbase.org/	The biggest online registry for miRNAs. Target results are provided by miRCosm, but no plant is supported.	[52]
	Rfam	√			http://rfam.sanger.ac.uk/	It is a collection of RNA families. miRNA family arrangement is different from miRBase.	[53]
Species-specific sources	ASRP	√	√		http://asrp.cgrb.oregonstate.edu/	*Arabidopsis* miRNAs and ta-siRNAs are collected, and also their targets.	[66]
	CSRDB	√	√	√	http://sundarlab.ucdavis.edu/smrnas/	Cereal sRNAs database, sRNAs of rice and maize are identified with 454 sequencing data.	[67]
miRNA annotation databases	MicroPC	√	√	√	http://www3a.biotec.or.th/micropc/index.html	A comprehensive resource for predicting and comparing plant miRNAs.	[54]
	PmiRKB	√	√		http://bis.zju.edu.cn/pmirkb/index.php	Four major functional modules are provided for plant miRNAs.	[55]
	PMRD	√	√		http://bioinformatics.cau.edu.cn/PMRD/	A plant-specific miRNA annotation database.	[56]
Target databases	miRTarBase		√		http://mirtarbase.mbc.nctu.edu.tw/index.html	Provides experimentally-verified miRNA-target interactions.	[60]
	starBase		√	√	http://starbase.sysu.edu.cn/index.php	Degradome-Seq data are used and five target prediction tools are integrated.	[61]
Genome & transcriptome databases	Phytozome				http://www.phytozome.net/	It provides 31 sequenced and annotated green plant genomes, which have been clustered into gene families at 11 evolutionarily-significant nodes.	[59]
	TAIR				http://www.arabidopsis.org/	TAIR maintains a database of genetic and molecular biology data for the model higher plant *Arabidopsis thaliana*.	[57]
	TIGR				http://rice.plantbiology.msu.edu/	TIGR provides genome sequence from the Nipponbare subspecies of rice and annotation of the 12 chromosomes.	[58]
High-throughput data	SRA				http://www.ncbi.nlm.nih.gov/sra	Sequence Read Archive (SRA) is a public repository for next-generation sequence data. sRNA sequencing data could be archived.	[64]
	MPSS				http://mpss.udel.edu/	Support several plant species, both sRNA and Degradome-Seq data are available.	[63]
	GEO				http://www.ncbi.nlm.nih.gov/geo/	Gene Expression Omnibus (GEO) database is a public repository for high-throughput gene expression data. Degradome-Seq data is also collected.	[65]

Note: [1]Query, allow user to brows pre-computed targets; [2]Predict, accept user submitted miRNA and/or mRNA sequences, and return predicted results.

employed extensively to validate products of RISC-mediated cleavage in plants [41].

High-throughput methods

To map RNA–RNA cleavage sites more comprehensively, Addo-Quaye et al. [74] and German et al. [75] introduced the parallel analysis of RNA ends (PARE), also known as Degradome-Seq or genome-wide mapping of uncapped transcripts (GMUCT). Degradome libraries are constructed by ligation of polyA-enriched RNA samples to a custom RNA adaptor containing a 3′ MmeI site, followed by reverse transcription, second-strand synthesis, MmeI digestion, ligation of a 3′ dsDNA adaptor, gel purification and PCR amplification (**Figure 2**). The power of coupling PARE and high-throughput sequencing has been demonstrated by identifying widespread mRNA cleavage events regulated by miRNAs in *Arabidopsis* [74–76], rice [77,78], grapevine [79] and limited cleavages in mammals [80–82]. Because extensive base pairing between miRNAs and mRNAs leading to direct RISC-mediated cleavage does not appear to be a major mechanism of miRNA activity

Figure 2 The PARE protocol

Parallel analysis of RNA ends (PARE) is also known as genome-wide mapping of uncapped transcripts (GMUCT) or Degradome-seq. PARE is a modified 5′-RACE with high-throughput deep sequencing methods. After cleavage, the downstream sequence of the target site will not be degraded. Thus, it would be possible to preserve the cleavage site by adding a 5′ adaptor. Furthermore, replacing the original long downstream sequence with a shorter subsequence (20 nt) and a new 3′ double strand DNA adaptor will make it realistic to enhance the performance by combining deep sequencing methods after purification and amplification.

in mammals, the use of PARE is most suited to plant systems, where it identifies the large subset of miRNA targets that are subject to direct cleavage [83].

Degradome data can be scrutinized to find evidence of cleaved sRNA targets without resorting to computational predictions. Here, the current task is to develop effective and efficient pipelines to make better use of these data. So far, three pipelines have been proposed to process degradome data.

CleaveLand is the first general pipeline for detecting fragments diagnostic of sRNA-mediated cleavage from degradome sequencing experiments [84]. To begin with, degradome sequences will be matched to structural RNAs using Oligomap [85] to filter out structural RNAs. The purified dataset is then mapped to the transcriptome by Oligomap. A 26 nt long 'query' mRNA subsequence is generated by extracting 13 nt long sequences upstream and downstream of the location of the 5′ end of a matching degradome sequence. All query sequences are aligned to sRNA sequences using the Needle program in the EMBOSS package [86]. Alignments are then scored according to a previously-described scheme [87] and those with scores not exceeding a certain threshold and having at least one degradome sequence start with the 10th nucleotide of complementarity are retained. In order to differentiate spurious results from real targets, the pipeline re-runs using randomly-shuffled sRNA sequences, which have dinucleotide and trinucleotide compositions consistent with those of the input transcriptome, to estimate signal-to-noise ratios. This pipeline has been considered in many subsequent studies, including the well-known web server, starBase [61], whose contribution has been verified [7].

The second Degradome-Seq based analysis pipeline is SeqTar [88], which allows more mismatches and critical mismatch or G:U wobble pair at the position 10 or 11. This method introduces two statistics; not only the alignment between sRNA and query sequence is measured but also the abundance of reads at the center of query sequence is measured too. SeqTar uses a modified Smith–Waterman algorithm to align sRNA to a query sequence, and BLASTN to align degradome sequences to transcriptome. Because loose rules are considered, especially at the cleavage position, this pipeline always predicts many more miRNA:target pairs than CleaveLand [88]. Additionally, SeqTar could predict a potential sRNA if an accumulation of reads is found at a specific position, named as a peak, on a target when no input sRNAs contributed to this accumulation.

Recently, Folkes et al. described a new, user-friendly, cross-platform degradome analysis tool, PAREsnip, which enables flexible and comprehensive high-throughput target analysis, allowing users to identify genome-wide networks of sRNA−target interactions resulting in transcript cleavage [89]. Similar to CleaveLand and SeqTar, the input for PAREsnip includes transcriptome, degradome and sRNAome data. In this method, another short read alignment method, PatMaN [90] is employed to map degradome sequences. When searching for sRNA sequences that could
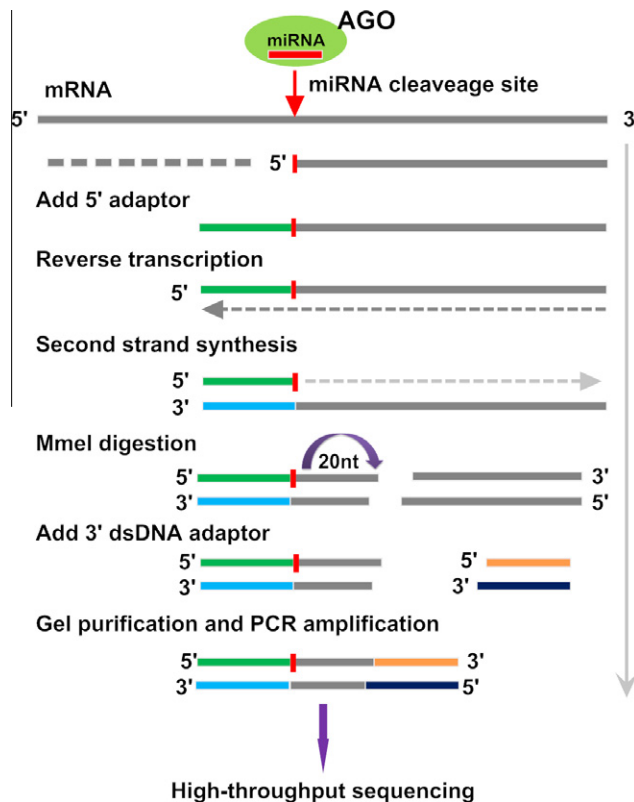
potentially cleave a transcript accounting for the degra-dome peak at a given position, the authors perform a pre-defined rule-based complementarity search algorithm [87,91], by traveling the 4-way tree [92]. sRNA−target interactions identified by PAREsnip will be further mea-sured by a *P*-value based metric, and only high quality pre-dictions will be retained. Although the complementary searching rules and selection score are similar, PAREsnip runs much faster than CleaveLand by performing the 4-way tree search. Additionally, PAREsnip employs cross-sample conservation of both sRNA and degradome sequences to reduce false predictions.

Another web tool, SoMART, is a collection of tools to process sRNAs [93]. Its slicer detector tool can predict sRNAs that could target the user-provided transcript, then the dRNA mapper tool can be used to align degradome sequences to the transcript. The server was recently updated with a new function, SMART COMPARE, to automatically compare the results from slicer detector and dRNA mapper, and identify slicers supported by degradome reads mapped to predicted cleavage sites.

Computational methods

As new experimental data comes to hand, the accepted modes of miRNA−target interaction are also expanded [94]. Although prediction programs may not incorporate all these experimentally-derived possibilities, their ability to provide potential targets easily and efficiently could greatly facilitate the downstream investigation. In this sec-tion, we will discuss the principles considered by existing computational methods. In order to facilitate description, we summarized popular available methods/tools in **Table 2** and annotated each method/tool based on principles men-tioned below.

Complementarity

For a miRNA, the complementarity between itself and its target site determines the stability of miRNA:target duplex and therefore has been utilized as a key feature for target gene analysis by computational methods.

Early observations suggested that ∼6 nt (position 2–7) *cis* elements that are required for post-transcriptional repression of *Drosophila melanogaster* targets are perfectly complementary to the 5′ ends of specific miRNAs [95]. Subsequently, systematic mutagenesis studies highlighted the seed regions for miRNA targeting in *Arabidopsis* [91,96]. The conserved pairing of the seed region markedly reduces the occurrence of false positive predictions. In gen-eral, a scoring schema that requires a perfect or nearly per-fect match within seed regions has been widely employed in published plant miRNA prediction methods/tools. How-ever, some exceptions have been reported. For example, miR398a post-transcriptionally regulates its target gene CSD2 in *Arabidopsis*, though the seed region contains a bulge and GU wobble [97]. Meanwhile, a set of canonical

seed types of different length or with specific initial base was verified by experiments [38], which puzzled the compu-tational methods. One recent study verified the specificity of long seeds but the majority of functional target sites are formed by less specific seeds of only 6 nt [98]. Addition-ally, they also found a substantial fraction of genuine tar-get sites are non-conserved. Chi et al. identified an alternative binding mode by which miR-124 can regulate its target with G-bugle site (positions 5–6), which cannot be explained by canonical seed matches [99]. All the contra-dictory evidence seems to require existing computational methods to review their seed region related scoring rules.

The degree of complementarity of the central region (position 9–11) is often presented as a decisive feature that determines whether slicing or translational repression and mRNA decay follows from RISC recruitment [100]. Addi-tionally, functional miRNA target sites that only pair with the central region of the miRNA lead to translational inhi-bition or mRNA decay in humans [80], which might be also present in plants. Thus, many computational methods con-sider this region when scoring a miRNA:target pair [101–104]. The differences are the position of the central region, and whether mismatch or wobble pair in this region is allowed, and if it is allowed whether the functional type of this miRNA:target pair is also predicted. As showed in Table 2, psRNATarget [101] and imiRTP [102] predict the functional type of miRNA based on the complemen-tary at the central region of the miRNA:target pair.

In addition to the seed and central regions, the miRNA 3′ backbone is also thought to be critical to enhance target rec-ognition in *Arabidopsis*. The Tcp4-soj18 mutation at posi-tion 19 of miR319 or position 16 of the mRNA target site significantly affects their pairing, although these changes cause only a small difference in the calculated interaction free energy [105]. In another study, Zhang et al. finds that perfect complementarity between the 3′ end of miR173 and the 5′ end of AT2G39675 (TAS1c) is crucial by system-atically mutating the miR173 target site [106]. Mismatches at 3′ end of miR173 abolish *trans*-acting siRNA (ta-siRNA) formation, while mismatches at the 5′ end had less effect. Unlike the seed and central region, the 3′ backbone was underrated by all existing computational methods.

The advantages and disadvantages of using different sets of complementarity are that considering only stringent-pair types increases specificity but might miss many potential targets, whereas considering both stringent and moder-ate-stringent-pair types increases sensitivity but might also increase the number of false positives.

Target site accessibility

The secondary structure is very important for predicting both miRNAs and miRNA targets [107,108]. An effective miRNA−target interaction begins with the hybridization reaction on an open structure at the target site (**Figure 3**). After binding, miRISC can disrupt the secondary structure at the site to elongate hybridization [109,110]. Kiryu et al.

Table 2 Plant miRNA target prediction tools

Name	Comp[1]	Cons[2]	Hyb[3]/Acc[4]	Mul[5]	Fun[6]	Availability[7]	Own sequence[8]		Link	Refs.
							miRNA	mRNA		
PatScan	•					L	yes	yes	N/A	[131]
miRNAassist	•			•		L[9]	yes	yes	N/A	[132]
miRU	•	•				W	no	no	N/A	[133]
WMD3	•		•			W	yes	no	http://wmd3.weigelworld.org/	[114]
TAPIR	•	•	•			W	yes	yes	http://bioinformatics.psb.ugent.be/webtools/tapir/	[115]
UEA sRNA	•		•			W	yes	no	http://srna-tools.cmp.uea.ac.uk/plant/	[103]
Target-align	•					L	yes	yes	http://www.leonxie.com/	[104]
Targetfinder	•	•				L	yes	yes	http://carringtonlab.org/resources/targetfinder	[19]
p-TAREF	•	•	•			W/L	yes	yes	http://scbb.ihbt.res.in/new/p-taref/form1.html	[116]
psRNATarget	•	•	•	•	•	W	yes	yes	http://plantgrn.noble.org/psRNATarget/	[101]
imiRTP	•	•	•	•	•	L	yes	yes	http://admis.fudan.edu.cn/projects/imiRTP.htm	[102]

Note: [1] Complementary; [2] Conservation; [3] Hybridization; [4] Accessibility; [5] Multiplicity; [6] Function; [7] W/L indicate whether the tool could be accessed online at website or installed locally; [8] yes/no indicate whether users' own miRNA and/or mRNA sequences could be used by the tool or not; [9] available upon request.

performed a detailed investigation of accessibility of target sites recently [111]. They found that the efficacy of miR-NAs depends strongly on the accessibility of both the very 5′ and 3′ end of their binding sites, which supports existing bioinformatics practice that extracts a certain length subsequence from upstream and downstream of target sites when computationally calculating the accessibility.

The free energy between a hybridization of miRNA and its target is considered by some computational methods, but it is not a good indicator, especially compared with the accessibility [112,113]. UEA sRNA [103], WMD3 [114], TAPIR [115] and p-TAREF [116] first calculate the hybridization energy of miRNA:target duplex, then compare it with the optimal hybridization energy, and finally use the percentage as a filter to choose potential targets. Another two methods, psRNATarget and imiRTP, calculate the accessibility by RNAup program in Vienna RNA package [117].

However, despite being theoretically sound, calculating accessibility could be extremely time-consuming, especially when the mRNA sequence is long. Moreover, current thermodynamic models used in RNA secondary structure prediction algorithms are not very accurate [118].

Evolutionary conservation of target sites

miRNA families are comprised to have the same seed site, and are well-conserved among related species [119]. In addition, miRNA families have targets that are conserved among related species. In early target prediction methods, exploring conserved miRNA targets in homology sequences within related species have been used to re-enforce the target prediction [120,121]. Several recent methods, like Targetfinder [19], TAPIR, p-TAREF, psRNATarget and imiRTP, still consider this principle as part of a scoring schema.

The main challenge of such a principle is that the expression pattern of miRNA targets may be time- and space-specific, thus target sequences may not be detected due to their low expression level in specific organisms [75], leading to false negative predictions from conservative analysis. On
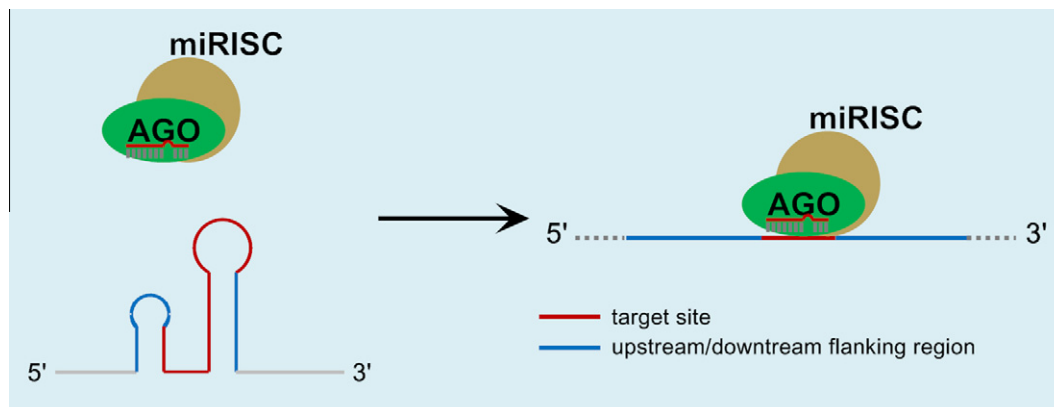
Figure 3 Target site accessibility
Accessibility of target mRNA and miRNA is believed to increase the precision of miRNA target predictions because the secondary structure around target site will prevent miRNA and mRNA target from contacting. Many methods mentioned in this article consider this fact by employing various tools (like RNAup, RNAduplex, *etc.*) to calculate target site accessibility, which is represented by the energy required to open secondary structure around target site. The less energy always means the more possibility that miRNA is able to contact target mRNA.
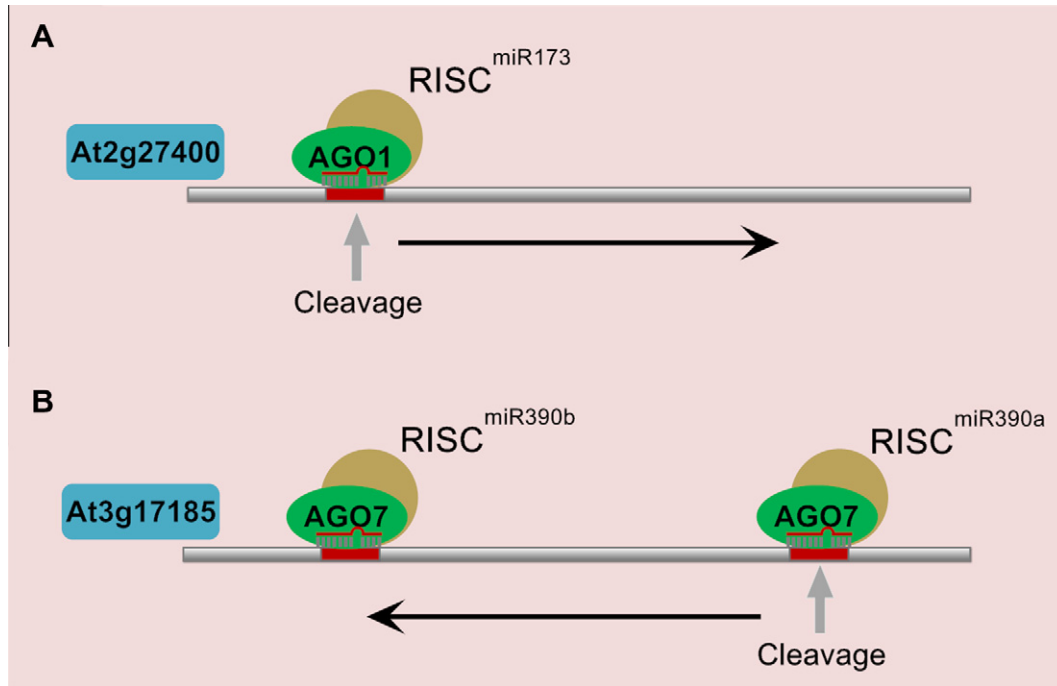
Figure 4 Target site multiplicity
In most situations, one miRNA is enough to change the expression of target genes. **A.** miR173 can trigger AtTAS1A (At2g27400) to generate ta-siRNAs. But there are some exceptions. **B.** Both miR390a and miR390b are crucial in the generation of ta-siRNAs from AtTAS3 (At3g17185).

top of that, there are also species-specific miRNAs and targets. Applying a conservation filter can decrease the false positive rate, but the side effect is also obvious, since it is only effective for conserved miRNAs. When species-specific miRNAs are of interest, it is important to identify targets both with and without conservation.

Besides the aforementioned principles, considering the number of putative miRNA sites per mRNA can significantly enhance target prediction [7]. Unfortunately, the importance of target site multiplicity was generally underestimated while highlighting the perfect match of miRNA and its target site (Table 2). One good example is the biogenesis of AtTAS3 to generate ta-siRNA (Figure 4). Axtell et al. firstly proposed the two-hit trigger model, that is, AtTAS3 is targeted by miR390a and miR390b simultaneously before giving rise to ta-siRNA [47]. Later, Montgomery et al. reported that though miRNA-guided cleavage only occurs at one miRNA binding site, the other miRNA binding site is still necessary for AtTAS3 processing [122]. In legumes, an AP2-like gene Medtr2g093060 contains a miR172-cleaved target site and a non-cleaving miR156 target site [123], which is another good example supporting the importance of multiplicity. Based on our latest work (Ding et al., in preparation), we speculate that this phenomenon is more common in plants.

Conclusion

The recent discovery that miRNAs can both regulate and be regulated by target interactions has profound implications for understanding their roles in gene regulation [124]. More questions about how miRNA targeting functions *in vivo* are raised after the identification of the intertwined relationship. An original and remaining challenge in the field is the ability to find miRNA targets with high confidence. Actually, finding true functional miRNA targets is still challenging even though many biological principles of miRNA targeting have been revealed experimentally and computationally.

A perfectly-complementary region on mRNA may not function as an effective target site due to its accessibility in terms of secondary structure and other unknown reasons, leading to false positive predictions [125]. Another problem that has hardly been addressed is multiplicity, since different miRNAs can cooperatively regulate individual targets, but miRNA expression signatures differ between cell types and conditions [27]. Thus, miRNA research will increasingly focus upon miRNA-regulated networks [126], in addition to identifying individual miRNA target interactions. Furthermore, besides repressing target gene expression, the reciprocal effect of targets on miRNAs is not entirely clear. In some cases, target interactions offer a protective influence on miRNA stability, whereas in others the outcome is miRNA degradation. Thus far, it seems that the degree of complementarity between a miRNA and its target determines the fate of the miRNA, and it seems that extensive complementarity favors miRNA destruction [127]. Although so many questions remain to be solved, recent progress in miRNA biology makes it clear that combining multiple strategies is required to obtain a comprehensive high-confidence description of miRNA targeting networks.

Integration of various data sources

Unlike miRNAs in animals, plant miRNAs predominately bind mRNAs at CDS, thus many researches focus on CDS while UTRs are ignored. However, new findings indicate that both animal and plant miRNAs can target 5′UTR, 3′UTR and coding regions [6,7], although the proportion that target UTRs is small. Expression level of target mRNAs or proteins is generally negatively associated with that of its corresponding miRNAs. The use of transcriptome data, particularly microarray and RNA-seq based expression data, should significantly reduce false positive predictions [128]. In addition, experimentally-verified miRNA: target pairs were collected by previous studies [115] and specific databases (like miRTarBase). Instead of using them as test data, p-TAREF originally runs over this information and applies it in the development of machine learning based methods and thus archives better results than other methods (more details, please see [116]).

Integration of known principles

None of the existing prediction tools has been able to incorporate all currently-known principles (Table 2), even those mentioned in this review. Directly building a model with more principles incorporated might provide higher accuracy and enhance site recognition efficacy, but its implementation might also become more complex. The appearance of imiRTP shed lights on further direction by integrating existing popular computational methods to get quality results, which is inspired by the successful experience of animal miRNA target study [102]. Prior to imiRTP, a series of such methods were proposed, but none of them support plant miRNAs. With the increase of novel target prediction methods, more species will be supported and more accurate results will be provided by such tools.

Integration of different technologies

New high-throughput technologies have accelerated the discovery of sequences that are bound by the miRNA complex *in vivo* [74,75]. These datasets provide an experimental framework for training computational algorithms to predict the likelihood of a sequence being recognized by specific miRNAs in a biologically-relevant context. Recently, several pipelines have been proposed for analysis of these high-throughput degradome-seq data and can scale miRNA targets at a genome-wide level. Remaining challenges for these methods come from several aspects. First, the computational cost is still high, including computation time and hardware cost. Second, current pipelines always generate much more candidate pairs than traditional computational methods, which is hard to believe that all of them are real predictions. Third, considering degradome-seq uses similar principle as 5′-RACE, novel experimental methods are needed to evaluate candidates identified by these pipelines. Fourth, with degradome-seq technology,

only cleavable targets can be identified, while non-cleavable targets will be missed. starBase provides targets predicted by integration of different computational methods and high-throughput technologies. How to use these data will be an interesting and meaningful question.

Along with the development of next-generation technology, hundreds of thousands of nucleic acid data are produced. In order to handle this scale of data, popular target prediction tools need to enhance their input/output ability as well as the ability of parallel calculating. For several tools listed in Table 2, we have tested their abilities (more details can be found in Table 1 in [7]). Thus, there will be a trend to involve compute unified device architecture (CUDA) [129] and cloud computing [130], which are newly-developed high-performance computing techniques, to meet the rapidly-increasing demand.

Competing interests

The authors declare no competing interests.

Acknowledgements

References

[1] Ghildiyal M, Zamore PD. Small silencing RNAs: an expanding universe. Nat Rev Genet 2009;10:94–108.

[2] Lee RC, Feinbaum RL, Ambros V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to lin-14. Cell 1993;75:843–54.

[3] Wightman B, Ha I, Ruvkun G. Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. Cell 1993;75:855–62.

[4] Reinhart BJ, Slack FJ, Basson M, Pasquinelli AE, Bettinger JC, Rougvie AE, et al. The 21-nucleotide *let-7* RNA regulates developmental timing in *Caenorhabditis elegans*. Nature 2000;403:901–6.

[5] Brennecke J, Stark A, Russell RB, Cohen SM. Principles of microRNA-target recognition. PLoS Biol 2005;3:e85.

[6] Lytle JR, Yario TA, Steitz JA. 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 2007;104:9667–72.

[7] Ding J, Li D, Ohler U, Guan J, Zhou S. Genome-wide search for miRNA-target interactions in *Arabidopsis thaliana* with an integrated approach. BMC Genomics 2012;13:S3.

[8] Grimson A, Farh KK, Johnston WK, Garrett-Engele P, Lim LP, Bartel DP. MicroRNA targeting specificity in mammals: determinants beyond seed pairing. Mol Cell 2007;27:91–105.

[9] Breakfield NW, Corcoran DL, Petricka JJ, Shen J, Sae-Seaw J, Rubio-Somoza I, et al. High-resolution experimental and computational profiling of tissue-specific known and novel miRNAs in *Arabidopsis*. Genome Res 2012;22:163–76.

[10] Williamson V, Kim A, Xie B, Omari McMichael G, Gao Y, Vladimirov V. Detecting miRNAs in deep-sequencing data: a software performance comparison and evaluation. Brief Bioinform 2012. http://dx.doi.org/10.1093/bib/bbs010.

[11] Licatalosi DD, Mele A, Fak JJ, Ule J, Kayikci M, Chi SW, et al. HITS-CLIP yields genome-wide insights into brain alternative RNA processing. Nature 2008;456:464–9.

[12] Darnell RB. HITS-CLIP: panoramic views of protein-RNA regulation in living cells. Wiley Interdiscip Rev RNA 2010;1:266–86.

[13] Hafner M, Landthaler M, Burger L, Khorshid M, Hausser J, Berninger P, et al. Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP. Cell 2010;141:129–41.

[14] Konig J, Zarnack K, Rot G, Curk T, Kayikci M, Zupan B, et al. ICLIP reveals the function of hnRNP particles in splicing at individual nucleotide resolution. Nat Struct Mol Biol 2010;17:909–15.

[15] Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004;116:281–97.

[16] Kim VN. MicroRNA biogenesis: coordinated cropping and dicing. Nat Rev Mol Cell Biol 2005;6:376–85.

[17] Carthew RW, Sontheimer EJ. Origins and mechanisms of miRNAs and siRNAs. Cell 2009;136:642–55.

[18] Kim VN, Han J, Siomi MC. Biogenesis of small RNAs in animals. Nat Rev Mol Cell Biol 2009;10:126–39.

[19] Fahlgren N, Howell MD, Kasschau KD, Chapman EJ, Sullivan CM, Cumbie JS, et al. High-throughput sequencing of *Arabidopsis* microRNAs: evidence for frequent birth and death of MIRNA genes. PLoS One 2007;2:e219.

[20] Cuperus JT, Montgomery TA, Fahlgren N, Burke RT, Townsend T, Sullivan CM, et al. Identification of MIR390a precursor processing-defective mutants in *Arabidopsis* by direct genome sequencing. Proc Natl Acad Sci U S A 2010;107:466–71.

[21] Song L, Axtell MJ, Fedoroff NV. RNA secondary structural determinants of miRNA precursor processing in *Arabidopsis*. Curr Biol 2010;20:37–41.

[22] Werner S, Wollmann H, Schneeberger K, Weigel D. Structure determinants for accurate processing of miR172a in *Arabidopsis thaliana*. Curr Biol 2010;20:42–8.

[23] Bologna NG, Mateos JL, Bresso EG, Palatnik JF. A loop-to-base processing mechanism underlies the biogenesis of plant microRNAs miR319 and miR159. EMBO J 2009;28:3646–56.

[24] Addo-Quaye C, Snyder JA, Park YB, Li YF, Sunkar R, Axtell MJ. Sliced microRNA targets and precise loop-first processing of MIR319 hairpins revealed by analysis of the *Physcomitrella patens* degradome. RNA 2009;15:2112–21.

[25] Zhang W, Gao S, Zhou X, Xia J, Chellappan P, Zhou X, et al. Multiple distinct small RNAs originate from the same microRNA precursors. Genome Biol 2010;11:R81.

[26] Park MY, Wu G, Gonzalez-Sulser A, Vaucheret H, Poethig RS. Nuclear processing and export of microRNAs in *Arabidopsis*. Proc Natl Acad Sci U S A 2005;102:3691–6.

[27] Voinnet O. Origin, biogenesis, and activity of plant microRNAs. Cell 2009;136:669–87.

[28] Axtell MJ, Westholm JO, Lai EC. Vive la difference: biogenesis and evolution of microRNAs in plants and animals. Genome Biol 2011;12:221.

[29] Yu B, Yang Z, Li J, Minakhina S, Yang M, Padgett RW, et al. Methylation as a crucial step in plant microRNA biogenesis. Science 2005;307:932–5.

[30] Li J, Yang Z, Yu B, Liu J, Chen X. Methylation protects miRNAs and siRNAs from a 3′-end uridylation activity in *Arabidopsis*. Curr Biol 2005;15:1501–7.

[31] Hutvagner G, Simard MJ. Argonaute proteins: key players in RNA silencing. Nat Rev Mol Cell Biol 2008;9:22–32.

[32] Qi Y, Denli AM, Hannon GJ. Biochemical specialization within *Arabidopsis* RNA silencing pathways. Mol Cell 2005;19:421–8.

[33] Vaucheret H. Plant ARGONAUTES. Trends Plant Sci 2008;13:350–8.

[34] Sempere LF, Cole CN, McPeek MA, Peterson KJ. The phylogenetic distribution of metazoan microRNAs: insights into evolutionary complexity and constraint. J Exp Zoolog B Mol Dev Evol 2006;306:575–88.

[35] Ruby JG, Jan C, Player C, Axtell MJ, Lee W, Nusbaum C, et al. Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. Cell 2006;127:1193–207.

[36] Floyd SK, Bowman JL. Gene regulation: ancient microRNA target sequences in plants. Nature 2004;428:485–6.

[37] Cuperus JT, Fahlgren N, Carrington JC. Evolution and functional diversification of MIRNA genes. Plant Cell 2011;23:431–42.

[38] Bartel DP. MicroRNAs: target recognition and regulatory functions. Cell 2009;136:215–33.

[39] Huntzinger E, Izaurralde E. Gene silencing by microRNAs: contributions of translational repression and mRNA decay. Nat Rev Genet 2011;12:99–110.

[40] Brodersen P, Sakvarelidze-Achard L, Bruun-Rasmussen M, Dunoyer P, Yamamoto YY, Sieburth L, et al. Widespread translational inhibition by plant miRNAs and siRNAs. Science 2008;320:1185–90.

[41] Llave C, Xie Z, Kasschau KD, Carrington JC. Cleavage of Scarecrow-like mRNA targets directed by a class of *Arabidopsis* miRNA. Science 2002;297:2053–6.

[42] Cullen BR. Viruses and microRNAs: RISCy interactions with serious consequences. Genes Dev 2011;25:1881–94.

[43] Orom UA, Nielsen FC, Lund AH. MicroRNA-10a binds the 5′UTR of ribosomal protein mRNAs and enhances their translation. Mol Cell 2008;30:460–71.

[44] Pasquinelli AE. MicroRNAs and their targets: recognition, regulation and an emerging reciprocal relationship. Nat Rev Genet 2012;13:271–82.

[45] Allen E, Howell MD. MiRNAs in the biogenesis of *trans*-acting siRNAs in higher plants. Semin Cell Dev Biol 2010;21:798–804.

[46] Manavella PA, Koenig D, Weigel D. Plant secondary siRNA production determined by microRNA-duplex structure. Proc Natl Acad Sci U S A 2012;109:2461–6.

[47] Axtell MJ, Jan C, Rajagopalan R, Bartel DP. A two-hit trigger for siRNA biogenesis in plants. Cell 2006;127:565–77.

[48] Montgomery TA, Yoo SJ, Fahlgren N, Gilbert SD, Howell MD, Sullivan CM, et al. AGO1-miR173 complex initiates phased siRNA formation in plants. Proc Natl Acad Sci U S A 2008;105:20055–62.

[49] Howell MD, Fahlgren N, Chapman EJ, Cumbie JS, Sullivan CM, Givan SA, et al. Genome-wide analysis of the RNA-DEPENDENT RNA POLYMERASE6/DICER-LIKE4 pathway in *Arabidopsis* reveals dependency on miRNA- and tasiRNA-directed targeting. Plant Cell 2007;19:926–42.

[50] Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet 2010;11:204–20.

[51] Wu L, Zhou H, Zhang Q, Zhang J, Ni F, Liu C, et al. DNA methylation mediated by a microRNA pathway. Mol Cell 2010;38:465–75.

[52] Kozomara A, Griffiths-Jones S. MiRbase: integrating microRNA annotation and deep-sequencing data. Nucleic Acids Res 2011;39:D152–7.

[53] Gardner PP, Daub J, Tate J, Moore BL, Osuch IH, Griffiths-Jones S, et al. Rfam: wikipedia, clans and the "decimal" release. Nucleic Acids Res 2011;39:D141–5.

[54] Mhuantong W, Wichadakul D. MicroPC (microPC): a comprehensive resource for predicting and comparing plant microRNAs. BMC Genomics 2009;10:366.

[55] Meng Y, Gou L, Chen D, Mao C, Jin Y, Wu P, et al. PmiRKB: a plant microRNA knowledge base. Nucleic Acids Res 2011;39:D181–7.

[56] Zhang Z, Yu J, Li D, Zhang Z, Liu F, Zhou X, et al. PMRD: plant microRNA database. Nucleic Acids Res 2010;38:D806–13.

[57] Lamesch P, Berardini TZ, Li D, Swarbreck D, Wilks C, Sasidharan R, et al. The *Arabidopsis* information resource (TAIR): improved gene annotation and new tools. Nucleic Acids Res 2012;40:D1202–10.

[58] Ouyang S, Zhu W, Hamilton J, Lin H, Campbell M, Childs K, et al. The TIGR rice genome annotation resource: improvements and new features. Nucleic Acids Res 2007;35:D883–7.

[59] Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, et al. Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res 2012;40:D1178–86.

[60] Hsu SD, Lin FM, Wu WY, Liang C, Huang WC, Chan WL, et al. MiRTarBase: a database curates experimentally validated microRNA-target interactions. Nucleic Acids Res 2011;39:D163–9.

[61] Yang JH, Li JH, Shao P, Zhou H, Chen YQ, Qu LH. StarBase: a database for exploring microRNA–mRNA interaction maps from argonaute CLIP-Seq and degradome-Seq data. Nucleic Acids Res 2011;39:D202–9.

[62] Mardis ER. Next-generation DNA sequencing methods. Annu Rev Genomics Hum Genet 2008;9:387–402.

[63] Nakano M, Nobuta K, Vemaraju K, Tej SS, Skogen JW, Meyers BC. Plant MPSS databases: signature-based transcriptional resources for analyses of mRNA and small RNA. Nucleic Acids Res 2006;34:D731–5.

[64] Kodama Y, Shumway M, Leinonen R. International Nucleotide Sequence Database Collaboration. The Sequence Read Archive: explosive growth of sequencing data. Nucleic Acids Res 2012;40:D54–6.

[65] Barrett T, Troup DB, Wilhite SE, Ledoux P, Evangelista C, Kim IF, et al. NCBI GEO: archive for functional genomics data sets-10 years on. Nucleic Acids Res 2011;39:D1005–10.

[66] Backman TW, Sullivan CM, Cumbie JS, Miller ZA, Chapman EJ, Fahlgren N, et al. Update of ASRP: the *Arabidopsis* small RNA project database. Nucleic Acids Res 2008;36:D982–5.

[67] Johnson C, Bowman L, Adai AT, Vance V, Sundaresan V. CSRDB: a small RNA integrated database and browser resource for cereals. Nucleic Acids Res 2007;35:D829–33.

[68] VanGuilder HD, Vrana KE, Freeman WM. Twenty-five years of quantitative PCR for gene expression analysis. Biotechniques 2008;44:619–26.

[69] Towbin H, Staehelin T, Gordon J. Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci U S A 1979;76:4350–4.

[70] Anon. Rapid amplification of 5′ complementary DNA ends (5′RACE). Nat Methods 2005;2:629–30.

[71] Thomson DW, Bracken CP, Goodall GJ. Experimental strategies for microRNA target identification. Nucleic Acids Res 2011;39:6845–53.

[72] Kuhn DE, Martin MM, Feldman DS, Terry AV Jr, Nuovo GJ, Elton TS. Experimental validation of miRNA targets. Methods 2008;44:47–54.

[73] Baek D, Villen J, Shin C, Camargo FD, Gygi SP, Bartel DP. The impact of microRNAs on protein output. Nature 2008;455:64–71.

[74] Addo-Quaye C, Eshoo TW, Bartel DP, Axtell MJ. Endogenous siRNA and miRNA targets identified by sequencing of the *Arabidopsis* degradome. Curr Biol 2008;18:758–62.

[75] German MA, Pillay M, Jeong DH, Hetawal A, Luo S, Janardhanan P, et al. Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. Nat Biotechnol 2008;26:941–6.

[76] German MA, Luo S, Schroth G, Meyers BC, Green PJ. Construction of parallel analysis of RNA ends (PARE) libraries for the study of cleaved miRNA targets and the RNA degradome. Nat Protoc 2009;4:356–62.

[77] Li YF, Zheng Y, Addo-Quaye C, Zhang L, Saini A, Jagadeeswaran G, et al. Transcriptome-wide identification of microRNA targets in rice. Plant J 2010;62:742–59.

[78] Wu L, Zhang Q, Zhou H, Ni F, Wu X, Qi Y. Rice microRNA effector complexes and targets. Plant Cell 2009;21:3421–35.

[79] Pantaleo V, Szittya G, Moxon S, Miozzi L, Moulton V, Dalmay T, et al. Identification of grapevine microRNAs and their targets using high-throughput sequencing and degradome analysis. Plant J 2010;62:960–76.

[80] Shin C, Nam JW, Farh KK, Chiang HR, Shkumatava A, Bartel DP. Expanding the microRNA targeting code: functional sites with centered pairing. Mol Cell 2010;38:789–802.

[81] Bracken CP, Szubert JM, Mercer TR, Dinger ME, Thomson DW, Mattick JS, et al. Global analysis of the mammalian RNA degradome reveals widespread miRNA-dependent and miRNA-independent endonucleolytic cleavage. Nucleic Acids Res 2011;39:5658–68.

[82] Karginov FV, Cheloufi S, Chong MM, Stark A, Smith AD, Hannon GJ. Diverse endonucleolytic cleavage sites in the mammalian transcriptome depend upon microRNAs, *Drosha*, and additional nucleases. Mol Cell 2010;38:781–8.

[83] Eckardt NA. Investigating translational repression by microRNAs in *Arabidopsis*. Plant Cell 2009;21:1624.

[84] Addo-Quaye C, Miller W, Axtell MJ. CleaveLand: a pipeline for using degradome data to find cleaved small RNA targets. Bioinformatics 2009;25:130–1.

[85] Berninger P, Gaidatzis D, van Nimwegen E, Zavolan M. Computational analysis of small RNA cloning data. Methods 2008;44:13–21.

[86] Rice P, Longden I, Bleasby A. EMBOSS: the European molecular biology open software suite. Trends Genet 2000;16:276–7.

[87] Allen E, Xie Z, Gustafson AM, Carrington JC. MicroRNA-directed phasing during *trans*-acting siRNA biogenesis in plants. Cell 2005;121:207–21.

[88] Zheng Y, Li YF, Sunkar R, Zhang W. SeqTar: an effective method for identifying microRNA guided cleavage sites from degradome of polyadenylated transcripts in plants. Nucleic Acids Res 2012;40:e28.

[89] Folkes L, Moxon S, Woolfenden HC, Stocks MB, Szittya G, Dalmay T, et al. PAREsnip: a tool for rapid genome-wide discovery of small RNA/target interactions evidenced through degradome sequencing. Nucleic Acids Res 2012;40:e103.

[90] Prufer K, Stenzel U, Dannemann M, Green RE, Lachmann M, Kelso J. PatMaN: rapid alignment of short sequences to large databases. Bioinformatics 2008;24:1530–1.

[91] Schwab R, Palatnik JF, Riester M, Schommer C, Schmid M, Weigel D. Specific effects of microRNAs on the plant transcriptome. Dev Cell 2005;8:517–27.

[92] Goodrich MT, Tamassia R. Data structures and algorithms in Java. 4th ed. New York: Wiley; 2006.

[93] Li F, Orban R, Baker B. SoMART: a web server for plant miRNA, tasiRNA and target gene analysis. Plant J 2012;70:891–901.

[94] Elefant N, Berger A, Shein H, Hofree M, Margalit H, Altuvia Y. RepTar: a database of predicted cellular targets of host and viral miRNAs. Nucleic Acids Res 2011;39:D188–94.

[95] Lai EC. Micro RNAs are complementary to 3′ UTR sequence motifs that mediate negative post-transcriptional regulation. Nat Genet 2002;30:363–4.

[96] Mallory AC, Reinhart BJ, Jones-Rhoades MW, Tang G, Zamore PD, Barton MK, et al. MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5′ region. EMBO J 2004;23:3356–64.

[97] Jones-Rhoades MW, Bartel DP. Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. Mol Cell 2004;14:787–99.

[98] Ellwanger DC, Buttner FA, Mewes HW, Stumpflen V. The sufficient minimal set of miRNA seed types. Bioinformatics 2011;27:1346–50.

[99] Chi SW, Hannon GJ, Darnell RB. An alternative mode of microRNA target recognition. Nat Struct Mol Biol 2012;19:321–7.

[100] Brodersen P, Voinnet O. Revisiting the principles of microRNA target recognition and mode of action. Nat Rev Mol Cell Biol 2009;10:141–8.

[101] Dai X, Zhao PX. PsRNATarget: a plant small RNA target analysis server. Nucleic Acids Res 2011;39:W155–9.

[102] Ding J, Yu S, Ohler U, Guan J, Zhou S. imiRTP: an integrated method to identifying miRNA-target interactions in *Arabidopsis*

thaliana. IEEE International Conference on Bioinformatics and Biomedicine. 2011, p. 100–104.

[103] Moxon S, Schwach F, Dalmay T, Maclean D, Studholme DJ, Moulton V. A toolkit for analysing large-scale plant small RNA datasets. Bioinformatics 2008;24:2252–3.

[104] Xie F, Zhang B. Target-align: a tool for plant microRNA target identification. Bioinformatics 2010;26:3002–3.

[105] Palatnik JF, Wollmann H, Schommer C, Schwab R, Boisbouvier J, Rodriguez R, et al. Sequence and expression differences underlie functional specialization of *Arabidopsis* microRNAs miR159 and miR319. Dev Cell 2007;13:115–25.

[106] Zhang C, Ng DW, Lu J, Chen ZJ. Roles of target site location and sequence complementarity in *trans*-acting siRNA formation in *Arabidopsis*. Plant J 2012;69:217–26.

[107] Bonnet E, Wuyts J, Rouze P, Van de Peer Y. Evidence that microRNA precursors, unlike other non-coding RNAs, have lower folding free energies than random sequences. Bioinformatics 2004;20:2911–7.

[108] Ding J, Zhou S, Guan J. MiRenSVM: towards better prediction of microRNA precursors using an ensemble SVM classifier with multi-loop features. BMC Bioinformatics 2010;11:S11.

[109] Long D, Lee R, Williams P, Chan CY, Ambros V, Ding Y. Potent effect of target structure on microRNA function. Nat Struct Mol Biol 2007;14:287–94.

[110] Kertesz M, Iovino N, Unnerstall U, Gaul U, Segal E. The role of site accessibility in microRNA target recognition. Nat Genet 2007;39:1278–84.

[111] Kiryu H, Terai G, Imamura O, Yoneyama H, Suzuki K, Asai K. A detailed investigation of accessibilities around target sites of siRNAs and miRNAs. Bioinformatics 2011;27:1788–97.

[112] Hausser J, Landthaler M, Jaskiewicz L, Gaidatzis D, Zavolan M. Relative contribution of sequence and structure features to the mRNA binding of Argonaute/EIF2C-miRNA complexes and the degradation of miRNA targets. Genome Res 2009;19:2009–20.

[113] Li X, Zhang YZ. Computational detection of microRNAs targeting transcription factor genes in *Arabidopsis thaliana*. Comput Biol Chem 2005;29:360–7.

[114] Ossowski S, Schwab R, Weigel D. Gene silencing in plants using artificial microRNAs and other small RNAs. Plant J 2008;53:674–90.

[115] Bonnet E, He Y, Billiau K, Van de Peer Y. TAPIR, a web server for the prediction of plant microRNA targets, including target mimics. Bioinformatics 2010;26:1566–8.

[116] Jha A, Shankar R. Employing machine learning for reliable miRNA target identification in plants. BMC Genomics 2011;12:636.

[117] Lorenz R, Bernhart SH, Honer Zu Siederdissen C, Tafer H, Flamm C, Stadler PF, et al. ViennaRNA Package 2.0. Algorithms Mol Biol 2011;6:26.

[118] Eddy SR. How do RNA folding algorithms work? Nat Biotechnol 2004;22:1457–8.

[119] Ding J, Zhou S, Guan J. MiRFam: an effective automatic miRNA classification method based on *n*-grams and a multiclass SVM. BMC Bioinformatics 2011;12:216.

[120] Ha M, Pang M, Agarwal V, Chen ZJ. Interspecies regulation of microRNAs and their targets. Biochim Biophys Acta 2008;1779:735–42.

[121] Rhoades MW, Reinhart BJ, Lim LP, Burge CB, Bartel B, Bartel DP. Prediction of plant microRNA targets. Cell 2002;110:513–20.

[122] Montgomery TA, Howell MD, Cuperus JT, Li D, Hansen JE, Alexander AL, et al. Specificity of ARGONAUTE7-miR390 inter-action and dual functionality in TAS3 *trans*-acting siRNA forma-tion. Cell 2008;133:128–41.

[123] Zhai J, Jeong DH, De Paoli E, Park S, Rosen BD, Li Y, et al. MicroRNAs as master regulators of the plant NB-LRR defense gene family via the production of phased*, trans*-acting siRNAs. Genes Dev 2011;25:2540–53.

[124] Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? Cell 2011;146:353–8.

[125] Dai X, Zhuang Z, Zhao PX. Computational analysis of miRNA targets in plants: current status and challenges. Brief Bioinform 2011;12:115–21.

[126] Peter ME. Targeting of mRNAs by multiple miRNAs: the next step. Oncogene 2010;29:2161–4.

[127] Ameres SL, Horwich MD, Hung JH, Xu J, Ghildiyal M, Weng Z, et al. Target RNA-directed trimming and tailing of small silencing RNAs. Science 2010;328:1534–9.

[128] Garber M, Grabherr MG, Guttman M, Trapnell C. Computational methods for transcriptome annotation and quantification using RNA-seq. Nat Methods 2011;8:469–77.

[129] Gálvez S, Díaz D, Hernández P, Esteban FJ, Caballero JA, Dorado G. Next-generation bioinformatics: using many-core processor architecture to develop a web service for sequence alignment. Bioinformatics 2010;26:683–6.

[130] Roy NC, Altermann E, Park ZA, McNabb WC. A comparison of analog and next-generation transcriptomic tools for mammalian studies. Brief Funct Genomics 2011;10:135–50.

[131] Dsouza M, Larsen N, Overbeek R. Searching for patterns in genomic data. Trends Genet 1997;13:497–8.

[132] Xie FL, Huang SQ, Guo K, Xiang AL, Zhu YY, Nie L, et al. Computational identification of novel microRNAs and targets in *Brassica napus*. FEBS Lett 2007;581:1464–74.

[133] Zhang Y. MiRU: an automated plant miRNA target prediction server. Nucleic Acids Res 2005;33:W701–4.