4010–4021 *Nucleic Acids Research, 2009, Vol. 37, No. 12* *Published online 05 May 2009*
doi:10.1093/nar/gkp272

Comprehensive prediction of novel microRNA targets in *Arabidopsis thaliana*

Leonardo Alves-Junior[1], Sandra Niemeier[1], Arne Hauenschild[2], Marc Rehmsmeier[2,*] and Thomas Merkle[1,*]

[1]Genome Research & RNA-based Regulation, Faculty of Biology and [2]Bioinformatics of Regulation, Center for Biotechnology (CeBiTec), Bielefeld University, D-33594 Bielefeld, Germany

Received May 19, 2008; Revised April 1, 2009; Accepted April 11, 2009

ABSTRACT

MicroRNAs (miRNAs) are 20–24 nt long endogenous non-coding RNAs that act as post-transcriptional regulators in metazoa and plants. Plant miRNA targets typically contain a single sequence motif with near-perfect complementarity to the miRNA. Here, we extended and applied the program RNAhybrid to identify novel miRNA targets in the complete annotated *Arabidopsis thaliana* transcriptome. RNAhybrid predicts the energetically most favorable miRNA:mRNA hybrids that are consistent with user-defined structural constraints. These were: (i) perfect base pairing of the duplex from nucleotide 8 to 12 counting from the 5′-end of the miRNA; (ii) loops with a maximum length of one nucleotide in either strand; (iii) bulges with no more than one nucleotide in size; and (iv) unpaired end overhangs not longer than two nucleotides. G:U base pairs are not treated as mismatches, but contribute less favorable to the overall free energy. The resulting hybrids were filtered according to their minimum free energy, resulting in an overall prediction of more than 600 novel miRNA targets. The specificity and signal-to-noise ratio of the prediction was assessed with either randomized miRNAs or randomized target sequences as negative controls. Our results are in line with recent observations that the majority of miRNA targets are not transcription factors.

INTRODUCTION

MicroRNAs (miRNAs) constitute a class of small non-coding RNA molecules that regulate gene expression at the post-transcriptional level (1). Since the first reports from plants in 2002 (2,3), it has become evident that miRNAs are crucial regulators of very diverse signaling networks, including the development of leaves (4,5), flowers (6–9), shoots and roots (10–12) and vascular tissue (13), also small RNA biogenesis and function (14,15), sensing nutrient stress (16–19), oxidative tolerance (20) and responses to phytohormones (21–24). In contrast to animals, the preferred mechanism of action of miRNAs in plants is the cleavage of target mRNAs by the RNA-induced silencing complex (RISC), guided by the miRNA (25). Plant miRNAs have also been reported to act by repressing translation (6) or by inducing methylation of DNA (26). The miRBase repository currently lists 187 miRNA genes for *Arabidopsis thaliana* (27). A major issue in miRNA research is to establish their functions and the functions of their targets. Prediction and experimental validation of target mRNAs provide the first essential steps towards this task. In plants, miRNAs typically show high sequence complementarity to a single sequence motif within the open reading frame of their target mRNA, the miRNA binding site. This feature forms the basis for the prediction of miRNA targets (19,28–33).

The first prediction of miRNA targets in Arabidopsis employed a pattern match algorithm that allowed a maximum of three mismatches between the miRNA and its complementary binding site on the target mRNA (28). No bulge was allowed, and G:U base pairs were considered as mismatches. Although this approach was very restrictive, several miRNA targets were identified. Many target mRNAs were members of the same gene family, and this characteristic was considered as supporting the prediction. In a refined prediction approach, a simple penalty scoring was employed (19). Each mismatch nucleotide in the miRNA:mRNA duplex was given the value 1.0, G:U base pairs and bulge nucleotides were given the value 0.5 and 2.0, respectively, and a maximum score of

*To whom correspondence should be addressed. Tel: +49 521 106 8723; Fax: +49 521 106 6423; Email: tmerkle@cebitec.uni-bielefeld.de
Correspondence may also be addressed to Marc Rehmsmeier. Tel: +43 1 79044 9920; Fax: +43 1 79044 9001; Email: marc.rehmsmeier@gmi.oeaw.ac.at
Present addresses:
Leonardo Alves-Junior, John Innes Centre, Dept. of Cell and Developmental Biology, Norwich Research Park, Colney, Norwich NR4 7UH, UK
Marc Rehmsmeier, GMI - Gregor Mendel Institute of Molecular Plant Biology GmbH, Dr. Bohr-Gasse 3, A-1030 Vienna, Austria

3.5 was allowed. In addition, conservation of miRNA complementarity between Arabidopsis and rice mRNAs was used as an additional positive selection factor. By relaxing the restrictions of the previous approach concerning the number of bulges and mismatches, many novel miRNA targets could be discovered. However, the criterion of conservation between distantly related species potentially leads to a high number of false-negative predictions due to a lack of gene conservation between different species. A similar approach was applied by Wang *et al*. (33).

A microarray analysis of plants overexpressing specific miRNAs combined with structural analyses of validated miRNA:mRNA hybrids resulted in a set of rules that seem to be necessary for functional miRNA:mRNA interactions (34). Generally, the pairing in the 5′ part of the miRNA appears to be most important for function, and only one mismatch was found in the regions corresponding to the nucleotides 2–12 of the miRNA, which includes the presumptive mRNA cleavage site opposite positions 10 and 11. In the 3′-end of the miRNA, a mismatch loop was shown to be tolerated if it contains no more than two nucleotides in either strand. On the other hand, a perfect match in this part can compensate for the presence of up to two mismatches in the 5′ part. The minimum free energy (mfe) of the duplex should be at least 72% of a perfect match calculated with the same miRNA, and the actual value should be −30 kcal/mol or below.

An approach based on the comparison of mfe values of miRNA:mRNA hybrids was developed by Rusinov *et al*. (35). In this implementation, the first six nucleotides of the miRNA were used for an initial sliding-window search in all Arabidopsis transcripts for six Watson–Crick matches, or five Watson–Crick matches plus one G:U base pair. When a hit was found, a sequence of 32 nt including the hit was extracted and a hybridization structure of the miRNA:mRNA duplex was calculated with an RNA folding program. Rusinov *et al*. (35) also implemented structural filters based on validated miRNA:mRNA hybrids.

Most Arabidopsis miRNAs known to date are not present in any other genome (33,36–38). As a consequence, target predictions for these miRNAs cannot make use of evolutionary conservation of target mRNAs. In order to predict miRNA targets independently of target mRNA conservation in other species, Xie *et al*. (39) analyzed a set of predicted and validated miRNA targets to define more specific parameters concerning the position of mismatches, the number of G:U base pairs, and the percentage of the mfe of the actual miRNA:mRNA hybrid compared with a perfect match hybrid. Using this approach, miRNA targets were predicted that are unique to Arabidopsis (36,38).

As major difference of our prediction approach to previous ones, we extended and applied the miRNA target prediction tool RNAhybrid with respect to plant specificities and predicted novel miRNA targets in the complete transcriptome of *A. thaliana*. RNAhybrid (40,41) predicts the energetically most favorable miRNA:mRNA hybrids that are consistent with user-defined structural constraints. The formation of G:U base pairs is allowed, and predictions do not depend on evolutionary conservation of miRNA targets. Using mfe as an additional filter, we predicted more than 600 novel miRNA targets, in addition to already predicted/validated targets. As a test for our prediction parameters, 10 targets were subjected to experimental validation, five of which were confirmed. Comparison of the gene ontology (GO) molecular function classes of our predicted targets with those of previously predicted/validated targets shows that miRNA target genes are not strongly over-represented in any specific functional class such as transcription factors. Finally, we performed miRNA target predictions for 12 additional plant species.

MATERIALS AND METHODS

Data acquisition

The dataset of *A. thaliana* candidate mRNAs used was the TAIR7_cdna_20070425 dataset, which includes all transcribed sequences according to the annotation release TAIR 7.0. These candidate mRNAs were obtained from the TAIR website (ftp://ftp.arabidopsis.org/Sequences/blast_datasets). The miRNA dataset was downloaded from miRBASE release 9.2 (http://microrna.sanger.ac.uk). This release contains 184 Arabidopsis miRNA genes, classified into 106 gene families.

Prediction of miRNA targets

Predictions of miRNA targets were done with the program RNAhybrid (40,41). Originally designed for the prediction of miRNA targets in animals and for general-purpose analyses of RNA hybridization, RNAhybrid had to be adapted to the specific requirements of plant miRNA target prediction. RNAhybrid performs *in silico* hybridizations between a miRNA and a possible target mRNA in a way that optimizes the free energy of the hybridization (41). G:U base pairs are not treated as mismatches, since they contribute to the overall free energy of RNA:RNA hybrids, albeit less favorably than standard base pairs. In our search for plant miRNA:mRNA hybrids, the following setting was applied: (i) perfect base pairing of the duplex from nucleotide 8 to 12 counting from the 5′-end of the miRNA, (ii) loops with a maximum length of 1 nt in either strand, (iii) bulges with no more than one nucleotide in size, and (iv) end overhangs not longer than two nucleotides. The 'seed' region from nucleotide 8 to 12 of the duplex includes the presumptive cleavage site in the target sequence. Finally, the minimum mfe value of the miRNA:mRNA hybrid was required to be 70% of the mfe calculated for a perfect match, in accordance with (31). Predictions done with this initial mfe cut-off were later filtered for a minimum mfe of 75%.

Assessing specificity/signal-to-noise ratio

To estimate the specificity and signal-to-noise ratio of our prediction, two methods were employed. For the first method, 10 randomized sequences for each miRNA (Supplementary Table S1) were generated with the program SHUFFLE from the HMMER package

(42; http://hmmer.janelia.org), maintaining the dinucleotide frequencies of the original miRNAs. The number of predicted targets for these randomly generated miRNAs was compared with the number of predicted targets for the authentic miRNAs (references in Supplementary Tables S2 and S8). Specificities and signal-to-noise ratios were calculated for each miRNA, and the averages were defined as specificity and signal-to-noise ratio of the entire prediction. More specifically, if TP is the number of true positives and FP the number of false positives, specificity is defined as $specificity = TP/(TP + FP)$ and signal-to-noise ratio as $signal\text{-}to\text{-}noise = (TP + FP)/FP$. These numbers are related by $specificity = 1 - 1/signal\text{-}to\text{-}noise$. Note that (TP + FP) is merely the number of all positively predicted miRNA:target relationships on authentic data (original miRNAs versus original mRNAs), and FP is estimated from randomized data.

In the second approach, a non-redundant dataset was created (TAIR7_nr) out of the TAIR7_cdna_20070425 dataset by using only one gene model per gene, and 10 randomized datasets of the same size (i.e. number of sequences and size of each sequence) and the same dinucleotide frequency as the TAIR7_nr were generated. Taking a set of 84 authentic non-redundant mature miRNA sequences (Supplementary Table S1), searches for miRNA binding sites were done with the TAIR7_nr dataset and with the 10 shuffled cohorts, and specificities and signal-to-noise ratios were calculated as above.

In addition, predictions were sorted into groups according to the mfes of miRNA:mRNA duplexes relative to perfect-match mfes. Six cut-offs were applied: 70, 72, 75, 77, 80 and 85%. These values were chosen on the basis of the distribution of the percentage of mfe values compared to the mfe of perfect match hybrids of miRNA:mRNA hybrids of validated miRNA targets (Figure 1). Specificity and signal-to-noise ratio was estimated for each group as described above. Sensitivities were calculated as the percentages of experimentally validated targets (Supplementary Table S2) that were identified within the target prediction for each mfe cut-off group.

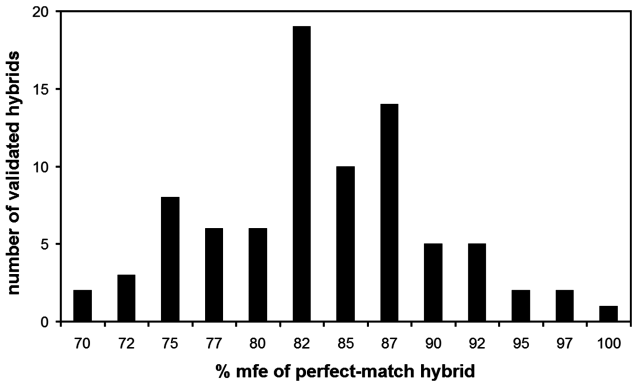

Figure 1. Distribution of mfe values of previously validated miRNA targets (Supplementary Table S2), given as percentage of a perfect-match hybrid, calculated with RNAhybrid.

Plant miRNA target database

Prediction results for *A. thaliana* and for other plant species are provided in a miRNA target database online (http://bibiserv.techfak.uni-bielefeld.de/cgi-bin/rnahybrid_tdb_mirnas.cgi). miRNA target predictions were performed with 12 additional plant species from which miRNAs have been isolated and registered in miRBase (http://microrna.sanger.ac.uk/). The transcript datasets were downloaded from the TIGR Plant Transcript Assemblies database (http://plantta.jcvi.org; 43).

Overexpression and detection of miRNAs

Sequences containing the precursors of ath-MIR156h, ath-MIR159a, ath-MIR161, ath-MIR172a, ath-MIR395b and ath-MIR414 (as listed in miRBase) were amplified by PCR with Phusion™ High-Fidelity DNA Polymerase (Finnzymes) from Col-0 genomic DNA as template in an Eppendorf Master Cycler using the primers listed in Supplementary Table S3. The resulting DNA fragments were ligated into the XbaI and SacI restriction sites of the pUC19-derived vector p35S 3′GFP (44), thereby replacing the cDNA encoding GFP. Each construct was used to transfect tobacco BY-2 (45) or Arabidopsis AT7 protoplasts (46), or both. Total RNA from transfected protoplasts was extracted using Tri-Reagent (Molecular Research Center) followed by DNaseI treatement (DNA-free™, Ambion). The expression of a given miRNA from the transgene was assayed by Northern blotting (2,47). A single-stranded RNA of 21 nucleotides in length (300 pMol/lane), separated along with the RNA samples, was used as a size marker. The choice of sequence allowed the detection with the same probe used to detect the U6snRNA (AT3G13855) for the gel loading control. In addition, for each miRNA, a DNA oligo with the same sequence of the miRNA was used as positive control for blot detection. The probes were prepared by labeling 20 μM of the specific DNA oligonucleotide with γ-[^{32}P]ATP (5000 Ci/mmol, 10 mCi/ml, Hartmann Analytic GmbH, Germany) using polynucleotide kinase (New England Biolabs) and purified with Sephadex G25 spin columns (GE Biosciences). The sequences of the probes are given in Supplementary Table S4.

Target validation

For experimental validation of selected miRNA targets, we used a modified protocol for RNA linker-mediated 5′ rapid amplification of cDNA ends (5′RACE), according to Llave *et al.* (48). The cDNAs of 10 miRNA targets, *ACS8, CKL6*, *GAE1*, *MRG1*, *MYB58, MYB94*, *MYB97*, *MYB101*, *MYB125* and *PRF2* (primers are listed in Supplementary Table S5), were amplified by PCR from the MatchMaker pGAD10 Arabidopsis cDNA library (Clontech), and ligated into the BamHI/XbaI and SmaI restriction sites of the pUC19-derived vector p35S 3′GFP (44), resulting in translational fusions with the cDNA encoding GFP. The expression of each GFP fusion protein was verified by transient transfection of BY-2 protoplasts and analysis of the GFP fluorescence by confocal

laser scanning microscopy using standard settings. To construct a 5′RACE library for each target, total RNA was isolated from AT7 protoplasts co-transfected with plasmids that overexpress both the miRNA precursor and the target mRNA. Alternatively, for target validation *in planta*, total RNA was extracted from young inflorescences of Arabidopsis Col-0 plants grown on soil. An RNA adaptor (300 ng, FirstChoice RLM-RACE kit, Ambion) was ligated to 10 μg of DNAseI treated total RNA using T4 RNA ligase (New England Biolabs) at 37°C for 1 h. For reverse transcriptase reactions, Superscript reverse transcriptase II (Invitrogen) was used according to the manufacturer's protocol. Reactions were done in 20 μl final volume with 2 μg of adaptor-ligated RNA, 1 mM oligo(dT) primer (dN-18dT), 500 μM of each dNTP, 1× First Strand Buffer, 10 μM of dithiothreitol and 20 U of reverse transcriptase II. The RNA adaptor provides an anchoring sequence for PCR primers. A PCR with the outer 5′RACE primer and a gene specific primer 1 (GSP1) was performed. A second nested PCR reaction was performed with the inner 5′RACE primer and a GSP2 primer (primers used for 5′RACE experiments are listed in Supplementary Table S6). PCR products were analyzed on a 1% agarose gel, and DNA fragments were directly ligated into TOPO-TA vector (TOPO-TA cloning kit, Invitrogen). Alternatively, PCR fragments were gel-purified using Qiaquick Gel Purification Kit (Qiagen) before ligation. Positive clones were screened by PCR using primers GSP3 and GSP1. For target mRNAs that contained the miRNA binding site at the very end of the open reading frame, gene specific primers for GFP were used. Between 5 and 10 positive clones were sequenced for each target.

RESULTS

Prediction of miRNA targets in Arabidopsis

The results of the calculations of specificities and sensitivities for six mfe cut-off groups are summarized in Table 1. A cut-off value of 75% of the mfe calculated for a perfect match hybrid with the same miRNA was chosen to filter the prediction due to high sensitivity (94.1%) and a signal-to-noise ratio of 4.5:1 and 5.2:1. We predicted 664 novel putative Arabidopsis miRNA targets. Previously predicted/validated targets that were also found with our prediction approach are not included in this number. The miRNA:mRNA hybrid structures of 26 examples of novel targets are shown in Figure 2. The complete list of novel miRNA targets of this prediction is presented in short in Supplementary Table S7 and can also be accessed in a plant miRNA target database online (http://bibiserv.techfak.uni-bielefeld.de/cgi-bin/rnahybrid_tdb_mirnas.cgi). Previously predicted/validated targets that were also predicted in this work are listed separately in Supplementary Table S8. For 12 additional plant species, miRNA target predictions were performed with the same parameters as described for Arabidopsis. The results are also included in the plant miRNA target database.

Novel predicted miRNA targets and gene ontology annotations of encoded proteins

For 84 miRNA families, novel miRNA targets were predicted, unequally distributed among the miRNA families (Supplementary Table S7). Over 48% of the novel targets were predicted for only nine miRNA families: miR396 (36 targets), miR413 (34 targets), miR773 (37 targets), miR781 (21 targets), miR834 (74 targets), miR837-5p (20 targets), miR838 (49 targets), miR847 (21 targets) and miR865-3p (57 targets). A total of 433 novel target candidates for miR414 were not included in these results because miR414 may not be a miRNA, and its expression is under debate (38,39).

Twenty-three miRNA families in the Arabidopsis genome are conserved in other plant genomes. For these miRNA families, many miRNA targets have been predicted and validated previously (19,28,33,39,49). Some conserved miRNAs target groups of similar genes, for example the miRNA families miR156/157, miR159 and miR167. The targets encode transcription factors of the Squamosa promoter-binding protein-like (SPL) group, MYB transcription factors (19,28) and auxin response factors (ARFs; 50), respectively. Among the novel predicted targets for conserved miRNAs, only a few are related to genes that are known to be 'miRNA-typical'. In our examples, two of five novel predicted miR156 targets and none of six novel miR157 targets encode SPL proteins, none of eight novel predicted miR159 targets code for MYB transcription factors, and none of six novel predicted miR167 targets encode ARFs (Supplementary Table S7). The majority of miRNA families is not

Table 1. Analysis of specificities/signal-to-noise ratios and sensitivities of the prediction

mfe cut-off (%)	Signal-to-noise (shuffled miRNAs)		Signal-to-noise (shuffled targets)		Sensitivity (%)
	Specificity (%)	Signal/noise	Specificity (%)	Signal/noise	
70	60	2.5	54.5	2.2	100
72	70.6	3.4	66.7	3.0	97.6
75	80.8	5.2	77.8	4.5	94.1
77	87	7.7	85.1	6.3	84.7
80	90.4	10.4	88.2	8.5	78.8
85	90.3	10.3	91.9	12.4	49.4

Mfe cut-offs were defined as the percentage of the mfe of the actual miRNA:mRNA hybrid compared to the mfe of a perfect match, calculated with RNAhybrid and with the same miRNA.

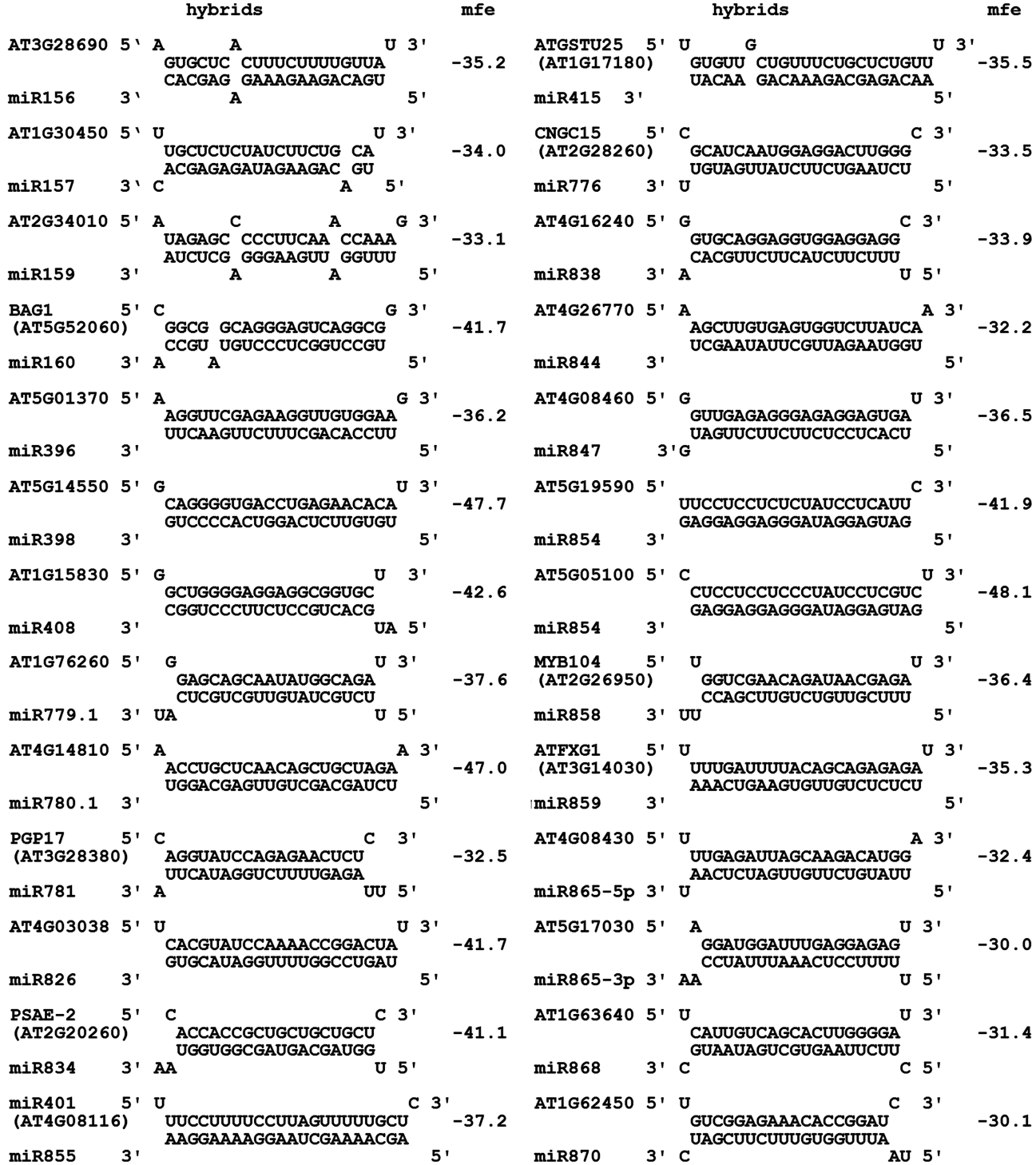

Figure 2. Calculated miRNA:mRNA hybrid structures for selected examples of novel miRNA targets from our prediction. Twenty-six structures of miRNA:mRNA hybrids predicted with RNAhybrid are presented. The miRNA binding site of the target mRNA is shown on top, the complementary miRNA as the bottom strand, calculated mfe values are given to the right (kcal/mol). AGI designation and, if applicable, the name of the gene as well as the designation of the miRNA (miRBase) are shown to the left.

conserved in other plant species and only found in Arabidopsis (32,33,36,38,51,52). Also for these miRNA families, many novel target mRNAs were predicted in this work, and most of them also encode proteins that belong to classes different to those predicted previously.

To gain more information on the novel miRNA targets that we predicted, GO annotations of molecular functions of the encoded proteins were extracted, and for each GO molecular function class the number of predicted targets was determined and compared to the categorization of the whole transcriptome (Figure 3). This analysis confirms

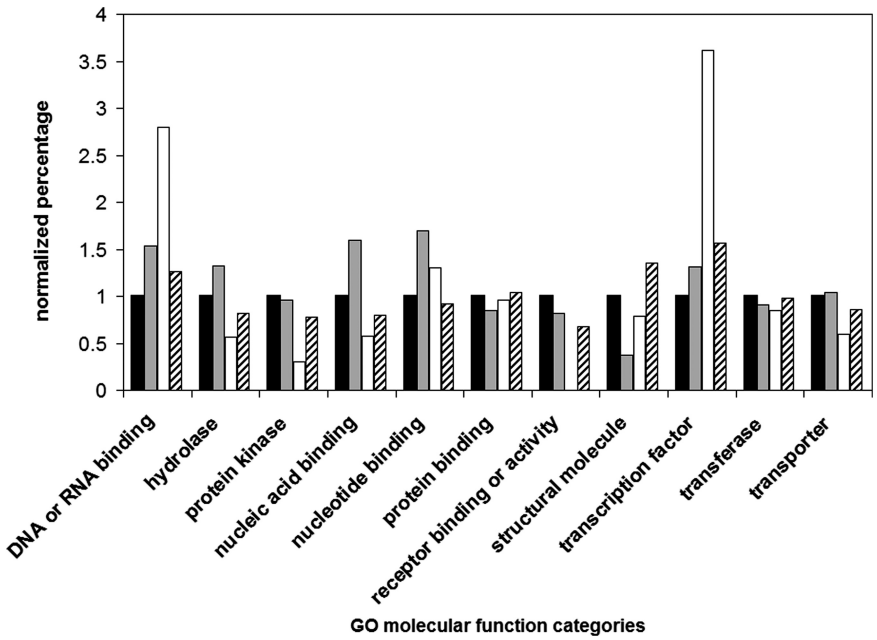

Figure 3. Analysis of gene ontology (GO) annotation terms for molecular function categories. The percentage of GO annotation terms for each category was normalized to the percentage of this category in the whole genome (black bars), which was set to 1. Novel predicted miRNA targets found with our approach are given as grey bars, previously predicted/validated targets are given as white bars. Hatched bars show the distribution among GO annotation terms of all miRNA targets predicted in this work, i.e. novel predicted and previously predicted/validated targets.

that targets are overrepresented in a small number of GO classes in the group of previously predicted/validated targets, namely transcription factors and DNA/RNA binding proteins, whereas targets are underrepresented in other GO classes, including transporters, kinases and hydrolases. In contrast, the complete dataset of our prediction including previously predicted/validated targets in addition to novel predicted targets does not show this highly unequal distribution between different GO classes. Although the GO classes of transcription factors and DNA/RNA binding proteins still contain more miRNA targets as compared to the whole transcriptome categorization, GO classes like kinases and hydrolases which had been highly underrepresented before, are now close to equal representation (Figure 3).

Validation of miRNA targets

We experimentally validated selected miRNA targets in AT7 or BY-2 protoplasts that transiently co-expressed the pre-miRNA and the target, as well as in Arabidopsis Col-0 plants. In the protoplast co-transfection assay, the expression of the target-GFP fusion protein and thus the production of the target RNA could be easily verified by monitoring the fluorescence (data not shown; 44,46). To verify the production of the mature miRNA in our assay system, Northern blots were prepared to detect small RNAs (Figure 4). Four out of six mature miRNAs assayed could be detected (miR156, miR159, miR161, miR395). For miR161 and miR395, a clearly higher expression level was detected in protoplasts transfected with the pre-miRNA than in protoplasts transfected with the empty vector. The miR395 was only detected in protoplasts transfected with the pre-miRNA. In contrast,

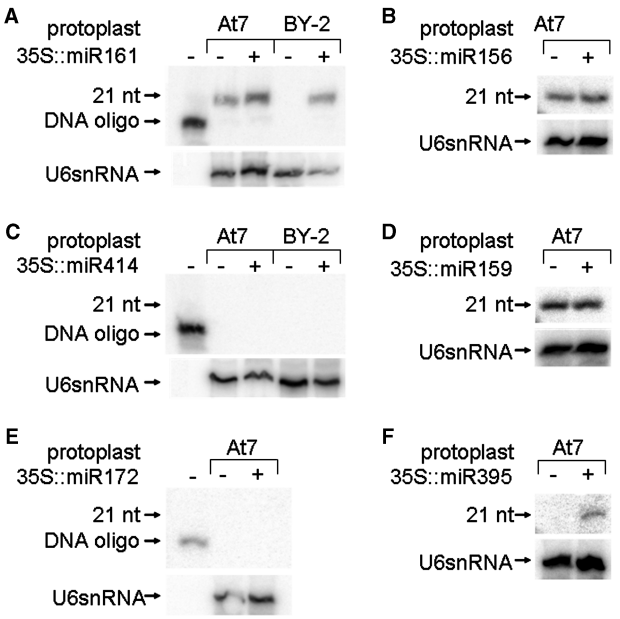

Figure 4. Detection of mature miRNAs in protoplasts. Protoplasts from Arabidopsis AT7 or tobacco BY-2 cell suspension cultures were transfected with plasmids harboring the precursors of (**A**) ath-MIR161, (**B**) ath-MIR156h, (**C**) ath-MIR414, (**D**) ath-MIR159a, (**E**) ath-MIR172a and (**F**) ath-MIR395b under the control of the 35S promoter. Total RNA was extracted and Northern blots to detect mature miRNAs were prepared from denaturating polyacrylamide gels. In each lane, 20 µg of total RNA was loaded from transfected (+) or untransfected (−) protoplasts. A positive control for transfer and hybridization, consisting of a DNA oligonucleotide with the same sequence as the corresponding mature miRNA, was included in all experiments (only shown in A, C and E). U6snRNA was used as loading control. An RNA oligonucleotide of 21 nucleotides in length was used as size marker. The position corresponding to 21 nucleotides is indicated.

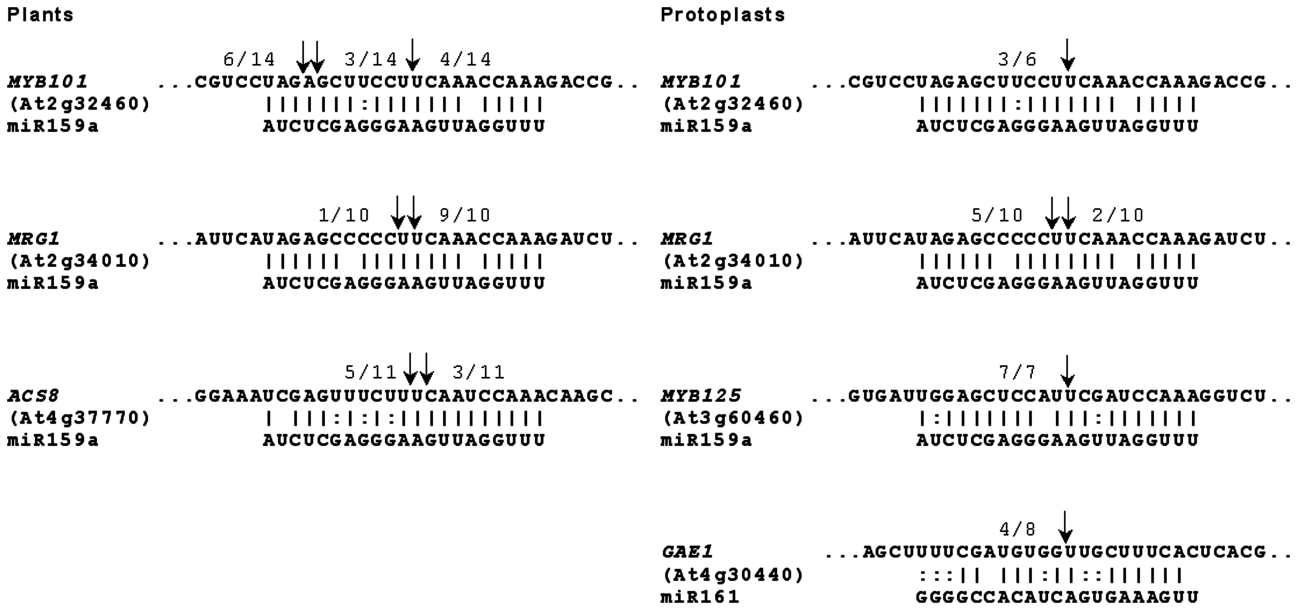

Figure 5. Five out of ten targets for four different miRNAs that were confirmed by validation experiments. *MYB101* and *MYB125* have been predicted previously (28). *MYB101*, *MRG1* and *ACS8* were predicted with our stringent parameters that we finally applied, *MYB125* and *GAE1* were predicted using less stringent parameters. During progress of our work, *MYB101*, *MYB125* and *MRG1* were also validated by others (24,31,55,56). Total RNA was extracted from young inflorescences of Arabidopsis wild-type plants or from AT7 protoplasts that were co-transfected with plasmids harboring the cDNA of the target and the precursor of the corresponding miRNA under the control of the 35S promoter. cDNAs were synthesized after ligation of an RNA linker to 5′RNA ends retaining a phosphate group. After amplification by two nested PCR reactions, DNA fragments were cloned and sequenced. Each panel shows part of the target mRNA (top) with the corresponding miRNA annealed to it (below). The name and Arabidopsis gene identifier for the target mRNA as well as the designation of the miRNA are given to the left. Watson–Crick base pairs are indicated with vertical dashes, G–U base pairs are indicated with a colon. Arrows indicate the cleavage sites and the corresponding relative numbers of analyzed 5′ RACE products.

although clearly detectable, there was no difference in the amount of miR156 and miR159 between pre-miRNA and mock transfections. Two miRNAs were not detected in Northern blot experiments, miR414 and miR172 (Figure 4C and E, respectively). We restricted subsequent validation experiments to targets of miR156, miR159, miR161 and miR395, since the expression of these miRNAs could be demonstrated.

We decided to validate ten targets for these four miRNAs that we predicted as a test for our prediction parameters. Three targets (*MYB101*, *MRG1*, *ACS8*) were predicted with the stringent parameters that we finally applied. Seven additional targets (*CKL6*, *GAE1*, *MYB58, MYB94*, *MYB97*, *MYB125* and *PRF2*) were predicted using less stringent parameters. Three targets (*MYB97*, *MYB101* and *MYB125*) had been predicted previously (19,28). None of these targets, however, were experimentally confirmed at the start of our experiments. Altogether, these experiments served as a test for the structural constraints of miRNA:mRNA hybrids and for the mfe cutoff that were applied in our final prediction. Using a 5′RACE strategy, we experimentally confirmed five targets: *MYB101*, as published by (24) during progress of our work, *MYB125*, also known as *DUO1* (53,54), *ACS8*, and *MRG1* (*miRNA-REGULATED GENE1*), all targets for miR159 (Figure 5). *MRG1* encodes a protein without conserved motifs and with no function assigned to date. *MYB125* and *MRG1* were also confirmed by others during progress of our work (55,56). We also confirmed *GAE1* (57), a miR161 target. In addition, using RNA extracted from Arabidopsis wildtype plants, we experimentally validated three miR159 targets by 5′RACE experiments: *MYB101*, *MRG1*, both with very similar results as in the protoplast system, and *ACS8* (Figure 5). The other five potential targets (*MYB94*, *PRF2*, *MYB58*, *MYB97* and *CKL6*) were not confirmed by 5′RACE (Figure 6). All of these five targets were predicted by us using less stringent parameters than those that were finally chosen. *MYB97* had been also predicted previously (19).

We performed database mining and looked for matches between our predicted targets and genes that are up-regulated in miRNA biogenesis mutants (29; Supplementary Tables S9, S10). We found 43 of the previously validated targets (out of Supplementary Table S8) and 83 of the novel predicted miRNA targets (out of Table S7). Both numbers are significantly higher than the number of matches expected by chance (Supplementary Table S10). We also compared our predictions with data from very recent approaches that identified ligation-competent mRNA 5′-ends/uncapped mRNAs or analyzed mutants that are affected in RNA metabolism and silencing, as summarized in Supplementary Table S11.

DISCUSSION

Prediction approach

A large number of different prediction approaches has been applied successfully in the identification of miRNA

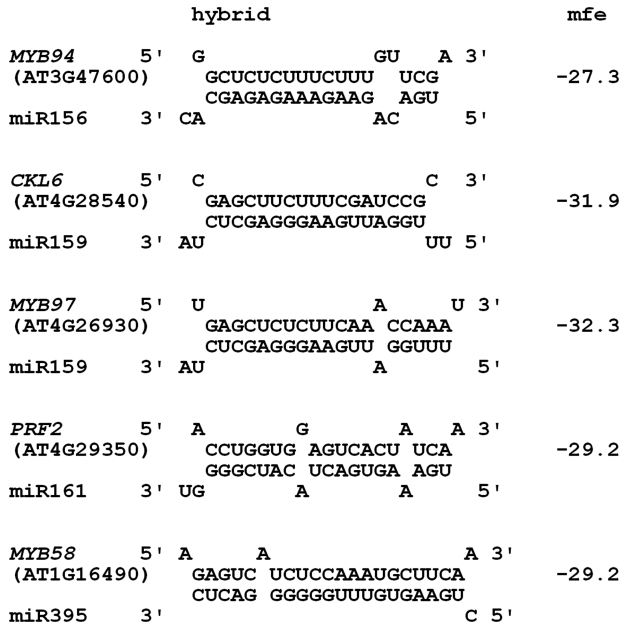

Figure 6. Five out of ten targets for four different miRNAs that were not confirmed by validation experiments. These targets were predicted using less stringent parameters than those that were finally applied. Therefore, they are not listed in the results of our final prediction. They served as experimental controls for the parameters of our final prediction. Experiments were done in parallel with those shown in Figure 5. MYB97 has been predicted previously (19). Hybrid structures calculated by RNAhybrid are shown, and the name and Arabidopsis gene identifier for the target mRNA as well as the designation of the miRNA are given to the left. The mfe (kcal/mol) of each hybrid is given to the right.

targets in plants (3,19,28,29,31–33,35,36,38,49,51,58). Naturally, most of the miRNA targets that show very high complementarity to their miRNAs were found first with rather stringent procedures that were based on very few parameters (28). More targets were predicted then with algorithms that allowed bulges (gaps) and/or loops (mismatches) and that considered conservation of target mRNAs in other plant species as an additional positive selection criterion (19). Many novel putative miRNA targets were found with our approach. As major difference to previous predictions, our approach is based on the calculation of miRNA:mRNA hybrids that are optimal with respect to their free energy. The program RNAhybrid (40,41) also allows the occurrence of G:U base pairs and does not treat them as mismatches. Parameters were chosen to accommodate structural features that have been found to be of importance in validated plant miRNA:mRNA hybrids. Perfect base pairing of the duplex from nucleotide 8 to 12 of the miRNA defined a 'seed' within the hybrid structure, reflecting the finding that mismatches are rare in the region around the target cleavage site. However, unpaired bases or mismatches are found frequently outside this seed in many miRNA: mRNA hydrids of validated plant targets. This was considered in our prediction, and loops and bulges were allowed with a maximum length of one nucleotide (on each strand for loops). Sequence parameters of plant and animal miRNA target recognition were reviewed and discussed very recently (59). Furthermore, we did not rely on any comparative genomics approach, since the majority of miRNAs found in Arabidopsis lack homologues in other plant species, and therefore this criterion cannot be used for target predictions in these cases (32,36,38,51). The fact that we nevertheless achieved high sensitivity and specificity values highlights the broad applicability of our approach. Altogether, these features resulted in the prediction of a large number of novel putative miRNA targets.

The results of our prediction are based on rather high stringency, since stretches of unpaired nucleotides ('gaps' or 'mismatches') in either miRNA or mRNA are restricted to one nucleotide in length. A relaxation of these parameters, and a change of the definition of the seed as well, results in a higher sensitivity but simultaneously also in a lower specificity. To further increase specificity we used an mfe cut-off, similar to (34). In estimating the specificity of our predictions, we defined six mfe cut-off groups and calculated signal-to-noise ratios and specificity values for each group (Figure 1, Table 1) based on predictions with real miRNAs (Supplementary Table S1) against randomized target datasets and with randomized miRNAs against the TAIR7 CDS dataset. Sensitivities of predictions were determined on the basis of experimentally validated Arabidopsis miRNA targets (Supplementary Table S2) as the percentage of identified targets of this dataset in the predictions for each mfe cut-off group. As summarized in Table 1, an mfe cutoff of 75% offered the best compromise with a high sensitivity of 94.1% and an acceptable signal-to-noise ratio between 4.5:1 and 5.2:1, corresponding to a specificity between 77% and 81%.

Novel predicted miRNA targets

Most of the previously predicted/validated miRNA targets are genes that encode transcription factors (1,60), and it was concluded that the high number of transcription factors among miRNA targets reflects the key role of miRNAs in gene regulatory networks (1). Analysis of the GO molecular function classes of proteins encoded by the novel miRNA targets predicted in this work including the previously predicted/validated targets showed that no major GO category was strongly overrepresented any more (Figure 3). Our work contributes to identifying novel putative targets among GO categories that were previously underrepresented, like transporter proteins, kinases and hydrolases. Thus, our data are in line with recent observations that the spectrum of miRNA-mediated regulation may be much broader than considered before. In plants, the number of targets per miRNA family is much smaller than in metazoa (61,62). Our work adds more than 600 novel predicted miRNA targets, and many of these shed new light on regulatory pathways that may be under miRNA control. This is illustrated by the discussion of a few selected examples.

The miR159 is an example of a miRNA controlling diverse biological processes (5,19,24,28,31). Among eleven previously predicted miR159 targets are seven genes that encode MYB transcription factors of the GAMYB group (63). In addition, *MYB125*/*DUO1* and

OPT1 were predicted as miR159 targets (31,55). *SPOROCYTELESS*, a putative novel target of miR159 predicted here, encodes a putative transcription factor that is involved in both micro- and megagametogenesis (64). Along with *MYB33* and *MYB65* that redundantly control anther development (65), and *MYB125*/*DUO1* that controls male gamete formation (54,57), *SPOROCYTELESS* is a novel target of miR159 that is also implicated in male fertility. On the other hand, *ACS8* (ACC SYNTHASE 8), reported as affected by miR159 overexpression (31), was predicted and confirmed as miR159 target in this work. Thus, miR159 regulates the expression of genes that are implicated in gibberellin signaling and ethylene biosynthesis.

For miR397, three genes encoding laccases were previously validated as targets (19). A novel predicted target, *DPA (DIMERIZATION PARTNER A)*, suggests that miR397 may also participate in the regulatory network that controls cell cycle. DPA is necessary for the function of E2F, a transcription factor that stimulates the transcription of genes necessary for G1-to-S and S phase progression during cell cycle (66,67).

miR319 is known to regulate the expression of *TCP* transcription factor genes whose down-regulations cause abnormalities in leaf development (5). miR319 may also regulate a gene involved in cell division and elongation in the growth zone of the root tip, *BREVIS RADIX* (*BRX*; 68), which has been predicted in this work.

miR156, miR159, miR164 and miR172 perform regulatory roles in flower development including the control of *LFY* expression, floral organ identity, and flowering time (6–8,21). miR413 may be another miRNA involved in flower development. Two of its targets predicted here affect the expression of *FLOWERING LOCUS C* (*FLC*) by means of chromatin modifications. FLC is a repressor protein that acts by inhibiting the floral transition (69). One of the predicted miR413 targets, *EARLY FLOWERING 8* (*ELF8*) is a gene encoding a protein that is required for histone 3 trimethylation at Lys 4 in the *FLC* chromatin. The reduced level of *FLC* chromatin methylation observed in *elf8* plants results in low expression of *FLC* and early flowering in both short and long-day conditions (70). The second putative target of miR413 that affects the *FLC* expression is *AtMBD9*, one among 13 Arabidopsis genes encoding proteins that contain a methyl-CpG-binding domain. In *atmbd9* plants, the early flowering phenotype is explained by the reduced level of *FLC* transcript, a consequence of a decreased level of acetylation in histones 3 and 4 of *FLC* chromatin (71). Thus, miR413 may regulate the level of *FLC* by two distinct mechanisms, although both mechanisms modify the state of *FLC* chromatin.

Target validation experiments

Experimental validation of target mRNAs in plant cells takes advantage of the preferred mode of action of plant miRNAs (72,73). In this work, 5′RACE experiments were performed with RNA samples extracted from Arabidopsis AT7 protoplasts co-transfected with plasmids to overexpress both the pre-miRNA and its putative target. The expression of the mature miRNA and the target mRNA (in fusion with GFP) in this system was tested by small RNA Northern blots (Figure 4) and analysis of the expression of GFP fluorescence, respectively. Alternatively, validation experiments were performed with RNA samples extracted from wild-type Arabidopsis plants. We decided to perform validation experiments with ten predicted targets for four different miRNAs. These targets were chosen to experimentally test our prediction parameters. Three targets (*MYB101, MRG1, ACS8*) were predicted with the stringent structural constraints and the mfe cutoff filter that we finally applied, the other targets were predicted with less stringent parameters (Figures 5 and 6). Out of these 10 targets, *MYB97*, *MYB101* and *MYB125* had been predicted previously by others (19,28). At the start of our work, however, none of these 10 targets had been experimentally confirmed. We successfully validated five targets (Figure 5), *MYB101*, *MYB125*, *MRG1, ACS8* (all targets for miR159) and *GAE1*, a target of miR161. In sum, we validated two previously predicted and three novel targets; *MYB101*, *MYB125* and *MRG1* were also validated by others during progress of our work (24,55,56). *ACS8* was reported to be affected by the over-expression of miR159 (31).

Five additional transcripts were not confirmed as miRNA targets: *MYB94* (miR156), *CKL6* (miR159), *MYB97* (miR159), *PRF2* (miR161) and *MYB58* (miR395). None of these transcripts nor the confirmed targets *GAE1* and *MYB125* are listed in Supplementary Table S7 or S8, since the miRNA:mRNA hybrids do not match the stringent structural constraints and the mfe cutoff filter that we finally applied. The miR161:*GAE1* hybrid does not satisfy the 75% mfe cutoff constraint, the miR159:*MYB125* and the miR161:*PRF2* hybrid show a mismatch within the seed at position 12, very close to the prospective cleavage site, and the miR156:*MYB94* hybrid has a mismatch loop containing two nucleotides on either strand (Figure 6). Nonetheless they were chosen for validation experiments because their hybrid structure showed less than five mismatches. The negative outcome of target validation experiments of the targets that did not match our stringent prediction parameters indicates that these parameters are effective to differentiate between true and false miRNA targets. However, the validation of *GAE1* as a miR161 target and *MYB125* as a target of miR159 confirms that some true miRNA targets may be among the candidates eliminated by these stringent criteria, as expected.

CONCLUSION

In this work, we report the prediction of many novel miRNA targets, suggesting that the repertoire of miRNA based regulation in plants is much broader than previously assumed. Many of our novel predictions extend the spectrum of biological processes that are subject to miRNA regulation. Furthermore, by changing the prediction parameters that we applied here accordingly, the Arabidopsis transcriptome may also be searched for

the presence of more mimicry targets like *IPS1* (74) by allowing 'mismatches' around the mRNA cleavage site. Our approach should be helpful in the prediction of more putative targets in other plant species, given the still increasing miRNA and transcriptome datasets.

Not all of these novel putative targets may be validated to be cleaved targets *in planta*. This may have several reasons. Some computationally predicted miRNAs that were not cloned or identified in deep sequencing studies may not be authentic miRNA genes (36–38). These include miR413, miR414, miR415 to miR420 and miR426 (33). On the other hand, there are still open questions concerning the structural requirements of miRNA:mRNA hybrid structures with respect to target selection and mode of function in plants (59). In addition, there is a higher degree of miRNA-based translational repression in plants than anticipated before (75). This may be a reason why only 43 of already known miRNA targets and 83 of our novel predicted targets were found to be up-regulated in plants carrying mutations in genes involved in miRNA biogenesis and function (29; Supplementary Table S9). Similarly, the expression levels of *MYB33* and *MYB65*, two known miR159 targets, were not reduced in plants overexpressing miR159 (31), which questions the suitability of microarray data as sole source to identify miRNA targets. A similar argument may hold true to explain the fact that only few of our predicted novel targets (Supplementary Table S10) were detected by very recent approaches to identify ligation-competent mRNA 5′-ends/uncapped mRNAs (56,76,77) or in mutants that are affected in RNA metabolism and silencing (78). With these findings in mind, *in silico* prediction continues to be a valuable tool for identifying potential plant miRNA targets.

SUPPLEMENTARY DATA

Supplementary data are available at NAR Online and Bielefeld University (http://bibiserv.techfak.uni-bielefeld.de/cgi-bin/rnahybrid_tdb_mirnas.cgi).

ACKNOWLEDGEMENTS

We are grateful to Bernd Weisshaar and Robert Giegerich for their continuous support. We thank Ute Bürstenbinder and Melanie Kuhlmann for excellent technical assistance, Prisca Viehöver and Angela Krahlemann for excellent sequencing service, and Agnes Bohne for help with validation experiments.

FUNDING

Deutsche Forschungsgemeinschaft (DFG, Bioinformatics Initiative BIZ 7/1 to M.R., T.M.). Funding for open access charge: OA Publication Fonds Bielefeld University. L.A.J., S.N. and A.H. were members of the International NRW Graduate School in Bioinformatics and Genome Research, and L.A.J. and S.N. received a fellowship from the DFG Research Training Group Bioinformatics (Graduiertenkolleg GK635).

Conflict of interest statement. None declared

REFERENCES

1. Jones-Rhoades,M.W., Bartel,D.P. and Bartel,B. (2006) MicroRNAS and their regulatory roles in plants. *Annu. Rev. Plant Biol.*, **57**, 19–53.
2. Llave,C., Kasschau,K.D., Rector,M.A. and Carrington,J.C. (2002) Endogenous and silencing-associated small RNAs in plants. *Plant Cell*, **14**, 1605–1619.
3. Reinhart,B.J., Weinstein,E.G., Rhoades,M.W., Bartel,B. and Bartel,D.P. (2002) MicroRNAs in plants. *Genes Dev.*, **16**, 1616–1626.
4. Emery,J.F., Floyd,S.K., Alvarez,J., Eshed,Y., Hawker,N.P., Izhaki,A., Baum,S.F. and Bowman,J.L. (2003) Radial patterning of Arabidopsis shoots by class III HD-ZIP and KANADI genes. *Curr. Biol.*, **13**, 1768–1774.
5. Palatnik,J.F., Allen,E., Wu,X., Schommer,C., Schwab,R., Carrington,J.C. and Weigel,D. (2003) Control of leaf morphogenesis by microRNAs. *Nature*, **425**, 257–263.
6. Aukerman,M.J. and Sakai,H. (2003) Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. *Plant Cell*, **15**, 2730–2741.
7. Baker,C.C., Sieber,P., Wellmer,F. and Meyerowitz,E.M. (2005) The early extra petals1 mutant uncovers a role for microRNA miR164c in regulating petal number in Arabidopsis. *Curr. Biol.*, **15**, 303–315.
8. Chen,X. (2004) A microRNA as a translational repressor of APETALA2 in Arabidopsis flower development. *Science*, **303**, 2022–2025.
9. Gandikota,M., Birkenbihl,R.P., Hohmann,S., Cardon,G.H., Saedler,H. and Huijser,P. (2007) The miRNA156/157 recognition element in the 3' UTR of the Arabidopsis SBP box gene SPL3 prevents early flowering by translational inhibition in seedlings. *Plant J.*, **49**, 683–693.
10. Guo,H.S., Xie,Q., Fei,J.F. and Chua,N.H. (2005) MicroRNA directs mRNA cleavage of the transcription factor NAC1 to downregulate auxin signals for arabidopsis lateral root development. *Plant Cell*, **17**, 1376–1386.
11. Laufs,P., Peaucelle,A., Morin,H. and Traas,J. (2004) MicroRNA regulation of the CUC genes is required for boundary size control in Arabidopsis meristems. *Development*, **131**, 4311–4322.
12. Mallory,A.C., Dugas,D.V., Bartel,D.P. and Bartel,B. (2004) MicroRNA regulation of NAC-domain targets is required for proper formation and separation of adjacent embryonic, vegetative, and floral organs. *Curr. Biol.*, **14**, 1035–1046.
13. Kim,J., Jung,J.H., Reyes,J.L., Kim,Y.S., Kim,S.Y., Chung,K.S., Kim,J.A., Lee,M., Lee,Y., Narry Kim,V. *et al.* (2005) microRNA-directed cleavage of ATHB15 mRNA regulates vascular development in Arabidopsis inflorescence stems. *Plant J.*, **42**, 84–94.
14. Vaucheret,H., Vazquez,F., Crete,P. and Bartel,D.P. (2004) The action of ARGONAUTE1 in the miRNA pathway and its regulation by the miRNA pathway are crucial for plant development. *Genes Dev.*, **18**, 1187–1197.
15. Xie,Z., Kasschau,K.D. and Carrington,J.C. (2003) Negative feedback regulation of Dicer-Like1 in Arabidopsis by microRNA-guided mRNA degradation. *Curr. Biol.*, **13**, 784–789.
16. Aung,K., Lin,S.I., Wu,C.C., Huang,Y.T., Su,C.L. and Chiou,T.J. (2006) pho2, a phosphate overaccumulator, is caused by a nonsense mutation in a microRNA399 target gene. *Plant Physiol.*, **141**, 1000–1011.
17. Chiou,T.J., Aung,K., Lin,S.I., Wu,C.C., Chiang,S.F. and Su,C.L. (2006) Regulation of phosphate homeostasis by MicroRNA in Arabidopsis. *Plant Cell*, **18**, 412–421.
18. Fujii,H., Chiou,T.J., Lin,S.I., Aung,K. and Zhu,J.K. (2005) A miRNA involved in phosphate-starvation response in Arabidopsis. *Curr. Biol.*, **15**, 2038–2043.
19. Jones-Rhoades,M.W. and Bartel,D.P. (2004) Computational identification of plant microRNAs and their targets, including a stress-induced miRNA. *Mol. Cell*, **14**, 787–799.
20. Sunkar,R., Kapoor,A. and Zhu,J.K. (2006) Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in Arabidopsis is mediated by downregulation of miR398 and important for oxidative stress tolerance. *Plant Cell*, **18**, 2051–2065.

21. Achard,P., Herr,A., Baulcombe,D.C. and Harberd,N.P. (2004) Modulation of floral development by a gibberellin-regulated microRNA. *Development*, **131**, 3357–3365.
22. Kasschau,K.D., Xie,Z., Allen,E., Llave,C., Chapman,E.J., Krizan,K.A. and Carrington,J.C. (2003) P1/HC-Pro, a viral suppressor of RNA silencing, interferes with Arabidopsis development and miRNA unction. *Dev. Cell*, **4**, 205–217.
23. Mallory,A.C., Bartel,D.P. and Bartel,B. (2005) MicroRNA-directed regulation of Arabidopsis AUXIN RESPONSE FACTOR17 is essential for proper development and modulates expression of early auxin response genes. *Plant Cell*, **17**, 1360–1375.
24. Reyes,J.L. and Chua,N.H. (2007) ABA induction of miR159 controls transcript levels of two MYB factors during Arabidopsis seed germination. *Plant J.*, **49**, 592–606.
25. Bartel,D.P. (2004) MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell*, **116**, 281–297.
26. Bao,N., Lye,K.W. and Barton,M.K. (2004) MicroRNA binding sites in Arabidopsis class III HD-ZIP mRNAs are required for methylation of the template chromosome. *Dev. Cell*, **7**, 653–662.
27. Griffiths-Jones,S. (2004) The microRNA Registry. *Nucleic Acids Res.*, **32**, D109–D111.
28. Rhoades,M.W., Reinhart,B.J., Lim,L.P., Burge,C.B., Bartel,B. and Bartel,D.P. (2002) Prediction of plant microRNA targets. *Cell*, **110**, 513–520.
29. Allen,E., Xie,Z., Gustafson,A.M. and Carrington,J.C. (2005) microRNA-directed phasing during trans-acting siRNA biogenesis in plants. *Cell*, **121**, 207–221.
30. Allen,E., Xie,Z., Gustafson,A.M., Sung,G.H., Spatafora,J.W. and Carrington,J.C. (2004) Evolution of microRNA genes by inverted duplication of target gene sequences in Arabidopsis thaliana. *Nat. Genet.*, **36**, 1282–1290.
31. Schwab,R., Palatnik,J.F., Riester,M., Schommer,C., Schmid,M. and Weigel,D. (2005) Specific effects of microRNAs on the plant transcriptome. *Dev. Cell*, **8**, 517–527.
32. Sunkar,R. and Zhu,J.K. (2004) Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. *Plant Cell*, **16**, 2001–2019.
33. Wang,X.J., Reyes,J.L., Chua,N.H. and Gaasterland,T. (2004) Prediction and identification of Arabidopsis thaliana microRNAs and their mRNA targets. *Genome Biol.*, **5**, R65.
34. Schwab,R., Ossowski,S., Riester,M., Warthmann,N. and Weigel,D. (2006) Highly specific gene silencing by artificial microRNAs in Arabidopsis. *Plant Cell*, **18**, 1121–1133.
35. Rusinov,V., Baev,V., Minkov,I.N. and Tabler,M. (2005) MicroInspector: a web tool for detection of miRNA binding sites in an RNA sequence. *Nucleic Acids Res.*, **33**, W696–W700.
36. Fahlgren,N., Howell,M.D., Kasschau,K.D., Chapman,E.J., Sullivan,C.M., Cumbie,J.S., Givan,S.A., Law,T.F., Grant,S.R., Dangl,J.L. *et al.* (2007) High-throughput sequencing of Arabidopsis microRNAs: evidence for frequent birth and death of miRNA genes. *PLoS ONE*, **2**, e219.
37. Lu,C., Tej,S.S., Luo,S., Haudenschild,C.D., Meyers,B.C. and Green,P.J. (2005) Elucidation of the small RNA component of the transcriptome. *Science*, **309**, 1567–1569.
38. Rajagopalan,R., Vaucheret,H., Trejo,J. and Bartel,D.P. (2006) A diverse and evolutionarily fluid set of microRNAs in Arabidopsis thaliana. *Genes Dev.*, **20**, 3407–3425.
39. Xie,Z., Allen,E., Fahlgren,N., Calamar,A., Givan,S.A. and Carrington,J.C. (2005) Expression of Arabidopsis MIRNA genes. *Plant Physiol.*, **138**, 2145–2154.
40. Rehmsmeier,M., Steffen,P., Hochsmann,M. and Giegerich,R. (2004) Fast and effective prediction of microRNA/target duplexes. *RNA*, **10**, 1507–1517.
41. Kruger,J. and Rehmsmeier,M. (2006) RNAhybrid: microRNA target prediction easy, fast and flexible. *Nucleic Acids Res.*, **34**, W451–W454.
42. Eddy,S.R. (2003) HMMER 2.3.2: biological sequence analysis using profile hidden Markov models.
43. Childs,K.L., Hamilton,J.P., Zhu,W., Ly,E., Cheung,F., Wu,H., Rabinowicz,P.D., Town,C.D., Buell,C.R. and Chan,A.P. (2007) The TIGR Plant Transcript Assemblies database. *Nucleic Acids Res.*, **35**, D846–D851.
44. Haasen,D., Kohler,C., Neuhaus,G. and Merkle,T. (1999) Nuclear export of proteins in plants: AtXPO1 is the export receptor for leucine-rich nuclear export signals in Arabidopsis thaliana. *Plant J.*, **20**, 695–705.
45. Merkle,T., Leclerc,D., Marshallsay,C. and Nagy,F. (1996) A plant in vitro system for the nuclear import of proteins. *Plant J.*, **10**, 1177–1186.
46. Hartmann,U., Valentine,W.J., Christie,J.M., Hays,J., Jenkins,G.I. and Weisshaar,B. (1998) Identification of UV/blue light-response elements in the Arabidopsis thaliana chalcone synthase promoter using a homologous protoplast transient expression system. *Plant Mol. Biol.*, **36**, 741–754.
47. Sambrook,J. and Russel,D.W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York.
48. Llave,C., Xie,Z., Kasschau,K.D. and Carrington,J.C. (2002) Cleavage of Scarecrow-like mRNA targets directed by a class of Arabidopsis miRNA. *Science*, **297**, 2053–2056.
49. Adai,A., Johnson,C., Mlotshwa,S., Archer-Evans,S., Manocha,V., Vance,V. and Sundaresan,V. (2005) Computational prediction of miRNAs in Arabidopsis thaliana. *Genome Res.*, **15**, 78–91.
50. Wu,M.F., Tian,Q. and Reed,J.W. (2006) Arabidopsis microRNA167 controls patterns of ARF6 and ARF8 expression, and regulates both female and male reproduction. *Development*, **133**, 4211–4218.
51. Lu,C., Kulkarni,K., Souret,F.F., MuthuValliappan,R., Tej,S.S., Poethig,R.S., Henderson,I.R., Jacobsen,S.E., Wang,W., Green,P.J. *et al.* (2006) MicroRNAs and other small RNAs enriched in the Arabidopsis RNA-dependent RNA polymerase-2 mutant. *Genome Res.*, **16**, 1276–1288.
52. Xie,Z., Allen,E., Wilken,A. and Carrington,J.C. (2005) DICER-LIKE 4 functions in trans-acting small interfering RNA biogenesis and vegetative phase change in Arabidopsis thaliana. *Proc. Natl Acad. Sci. USA*, **102**, 12984–12989.
53. Rotman,N., Durbarry,A., Wardle,A., Yang,W.C., Chaboud,A., Faure,J.E., Berger,F. and Twell,D. (2005) A novel class of MYB factors controls sperm-cell formation in plants. *Curr. Biol.*, **15**, 244–248.
54. Durbarry,A., Vizir,I. and Twell,D. (2005) Male germ line development in Arabidopsis. duo pollen mutants reveal gameto-phytic regulators of generative cell cycle progression. *Plant Physiol.*, **137**, 297–307.
55. Palatnik,J.F., Wollmann,H., Schommer,C., Schwab,R., Boisbouvier,J., Rodriguez,R., Warthmann,N., Allen,E., Dezulian,T., Huson,D. *et al.* (2007) Sequence and expression differences underlie functional specialization of arabidopsis microRNAs miR159 and miR319. *Dev. Cell*, **13**, 115–125.
56. German,M.A., Pillay,M., Jeong,D.-H., Hetawal,A., Luo,S., Janardhanan,P., Kannan,V., Rymarquis,L.A., Nobuta,K., German,R. *et al.* (2008) Global identification of microRNA-target RNA pairs by parallel analysis of RNA ends. *Nat. Biotech.*, **8**, 941–946.
57. Molhoj,M., Verma,R. and Reiter,W.D. (2004) The biosynthesis of D-Galacturonate in plants. Functional cloning and characterization of a membrane-anchored UDP-D-Glucuronate 4-epimerase from Arabidopsis. *Plant Physiol.*, **135**, 1221–1230.
58. Park,W., Li,J., Song,R., Messing,J. and Chen,X. (2002) CARPEL FACTORY, a Dicer homolog, and HEN1, a novel protein, act in microRNA metabolism in Arabidopsis thaliana. *Curr. Biol.*, **12**, 1484–1495.
59. Brodersen,P. and Voinnet,O. (2009) Revisiting the principles of microRNA target recognition and mode of action. *Nat. Rev.*, **10**, 141–148.
60. Dugas,D.V. and Bartel,B. (2004) MicroRNA regulation of gene expression in plants. *Curr. Opin. Plant Biol.*, **7**, 512–520.
61. Mallory,A.C. and Vaucheret,H. (2006) Functions of microRNAs and related small RNAs in plants. *Nat. Genet.*, **38 (Suppl 1)**, S31–S36.
62. Zhang,B., Wang,Q. and Pan,X. (2007) MicroRNAs and their regulatory roles in animals and plants. *J. Cell Physiol.*, **210**, 279–289.
63. Stracke,R., Werber,M. and Weisshaar,B. (2001) The R2R3-MYB gene family in Arabidopsis thaliana. *Curr. Opin. Plant Biol.*, **4**, 447–456.
64. Yang,W.C., Ye,D., Xu,J. and Sundaresan,V. (1999) The SPOROCYTELESS gene of Arabidopsis is required for initiation

of sporogenesis and encodes a novel nuclear protein. *Genes Dev*., **13**, 2108–2117.
65. Millar,A.A. and Gubler,F. (2005) The Arabidopsis GAMYB-like genes, MYB33 and MYB65, are microRNA-regulated genes that redundantly facilitate anther development. *Plant Cell*, **17**, 705–721.
66. Magyar,Z., De Veylder,L., Atanassova,A., Bako,L., Inze,D. and Bogre,L. (2005) The role of the Arabidopsis E2FB transcription factor in regulating auxin-dependent cell division. *Plant Cell*, **17**, 2527–2541.
67. Vandepoele,K., Raes,J., De Veylder,L., Rouze,P., Rombauts,S. and Inze,D. (2002) Genome-wide analysis of core cell cycle genes in Arabidopsis. *Plant Cell*, **14**, 903–916.
68. Mouchel,C.F., Osmont,K.S. and Hardtke,C.S. (2006) BRX mediates feedback between brassinosteroid levels and auxin signalling in root growth. *Nature*, **443**, 458–461.
69. Michaels,S.D. and Amasino,R.M. (1999) FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell*, **11**, 949–956.
70. He,Y., Doyle,M.R. and Amasino,R.M. (2004) PAF1-complex-mediated histone methylation of FLOWERING LOCUS C chromatin is required for the vernalization-responsive, winter-annual habit in Arabidopsis. *Genes Dev*., **18**, 2774–2784.
71. Peng,M., Cui,Y., Bi,Y.M. and Rothstein,S.J. (2006) AtMBD9: a protein with a methyl-CpG-binding domain regulates flowering time and shoot branching in Arabidopsis. *Plant J*., **46**, 282–296.
72. Chen,X. (2005) MicroRNA biogenesis and function in plants. *FEBS Lett*., **579**, 5923–5931.
73. Mallory,A.C., Reinhart,B.J., Jones-Rhoades,M.W., Tang,G., Zamore,P.D., Barton,M.K. and Bartel,D.P. (2004) MicroRNA control of PHABULOSA in leaf development: importance of pairing to the microRNA 5′ region. *EMBO J*., **23**, 3356–3364.
74. Franco-Zorrilla,J.M., Valli,A., Todesco,M., Mateos,I., Puga,M.I., Rubio-Somoza,I., Leyva,A., Weigel,D., Garcia,J.A. and Paz-Ares,J. (2007) Target mimicry provides a new mechanism for regulation of microRNA activity. *Nat. Genet*., **39**, 1033–1037.
75. Brodersen,P., Sakvarelidze-Achard,L., Bruun-Rasmussen,M., Dunoyer,P., Yamamoto,Y.Y., Sieburth,L. and Voinnet,O. (2008) Widespread translational inhibition by plant miRNAs and siRNAs. *Science*, **320**, 1185–1190.
76. Addo-Quaye,C., Eshoo,T.W., Bartel,D.P. and Axtell,M.J. (2008) Endogenous siRNA and miRNA targets identified by sequencing of the Arabidopsis degradome. *Curr. Biol*., **18**, 758–762.
77. Jiao,Y., Riechmann,J.L. and Meyerowitz,E.M. (2008) Transcriptome-wide analysis of uncapped mRNAs in Arabidopsis reveals regulation of mRNA degradation. *Plant Cell*, **20**, 2571–2585.
78. Gregory,B.D., O'Malley,R.C., Lister,R., Urich,M.A., Tonti-Filippini,J., Chen,H., Millar,A.H. and Ecker,J.R. (2008) A link between RNA metabolism and silencing affecting Arabidopsis development. *Dev. Cell*, **14**, 854–866.